

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 12:36:13 ; Search time 87.81 Seconds
(without alignments)
5591.862 Million cell updates/sec

Title: US-08-961-083-1

Perfect score: 1999

Sequence: 1 TAAATCTACGACAATAAA.....ATCTCAACAGCACAACCA 1999

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	3	US-08-961-083-1
2	1981.4	99.1	2049	3	US-08-481-435-5
3	944.2	47.2	960	2	US-08-245-511-3
4	944.2	47.2	960	2	US-08-600-993A-3
5	117.2	5.9	9100	2	US-08-743-637B-27
6	117.2	5.9	9100	3	US-08-526-840B-27
7	85	4.3	2487	3	US-08-481-435-1
8	62.2	3.1	807	2	US-08-771-716-1
9	62.2	3.1	807	2	US-08-771-716-3
10	62.2	3.1	807	3	US-09-057-720A-1
11	62.2	3.1	807	3	US-09-057-720A-3
12	56.4	2.8	2193	1	US-08-731-716-1
13	56.4	2.8	2193	1	US-08-731-716-3
14	55.2	2.8	428	4	US-09-060-756-206
15	53.8	2.7	1474	4	US-09-508-542-17
16	51.8	2.6	363	4	US-09-060-756-626
17	49.8	2.5	4403765	4	US-09-103-840A-2
18	49.8	2.5	4411529	4	US-09-103-840A-1
19	45.2	2.3	5511	3	US-08-928-361B-2
20	45.2	2.3	7334	3	US-08-928-361B-1
21	41.8	2.1	5163	3	US-08-700-651-1
22	41.8	2.1	5163	3	US-08-928-361B-4
23	41.8	2.1	5318	3	US-08-700-651-2
24	41.8	2.1	5318	3	US-08-928-361B-3
25	40.6	2.0	223	4	US-09-060-756-438
26	36.6	1.8	1172	4	US-08-936-165A-8
27	36.6	1.8	7218	1	US-08-232-463-14

```

c 28 36.4 1.8 473 4 US-09-060-756-671 Sequence 671, Appl
c 29 35.8 1.8 348 4 US-09-060-756-99 Sequence 99, Appl
c 30 35.2 1.8 43795 3 US-08-742-185-101 Sequence 101, Appl
c 31 34 1.7 2946 4 US-09-175-928-3 Sequence 3, Appl
c 32 33.8 1.7 1007 4 US-08-858-207A-113 Sequence 113, Appl
c 33 33.4 1.7 198 5 PCT-US95-10668-3 Sequence 3, Appl
c 34 33.4 1.7 198 5 PCT-US95-10668-4 Sequence 4, Appl
c 35 33.4 1.7 2029 4 US-09-136-574A-46 Sequence 46, Appl
c 36 33.2 1.7 700 4 US-09-236-097-11 Sequence 11, Appl
c 37 33 1.7 700 4 US-09-236-097-8 Sequence 8, Appl
c 38 32.8 1.6 366 4 US-09-060-756-331 Sequence 331, Appl
c 39 32.8 1.6 1725 1 US-08-257-073-14 Sequence 14, Appl
c 40 32.6 1.6 1543 3 US-08-714-918-82 Sequence 82, Appl
c 41 32.6 1.6 1543 4 US-09-265-315-82 Sequence 82, Appl
c 42 32.6 1.6 1543 4 US-09-265-315-82 Sequence 82, Appl
c 43 32.6 1.6 1543 4 US-09-266-417-82 Sequence 82, Appl
c 44 32.6 1.6 3343 4 US-08-965-762-21 Sequence 21, Appl
c 45 32.4 1.6 3292 1 US-07-814-964-12 Sequence 12, Appl

```

ALIGNMENTS

```

RESULT 1
US-08-961-083-1
; Sequence 1, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-083-1

```

Query Match 100.0%; Score 1999; DB 3; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Gaps 0;

QY 1 TAAATCTACGACAATAAAATCAACTGCTGACTTGGTTCGTAACGCCGCTCAA 60

DB 1 TAAATCTACGACAATAAAATCAACTGCTGACTTGGTTCGTAACGCCGCTCAA 60

QY 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
QY 121 TCGCTTCTTCGACACAGGGGATTTGATACCAATCGATTCCTGGGAGCTTTCTTCGGCAA 180
Db 121 TCGCTTCTTCGACACAGGGGATTTGATACCAATCGATTCCTGGGAGCTTTCTTCGGCAA 180
QY 181 TCGTCAAAAGCAATCCCTCCCAAGTGGATCAACTCCACCAACAGTTGATTAAGTTGAC 240
Db 181 TCGTCAAAAGCAATCCCTCCCAAGTGGATCAACTCCACCAACAGTTGATTAAGTTGAC 240
QY 241 TTACTTTTCAACTTCGACTTCGACAGCAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT 300
Db 241 TTACTTTTCAACTTCGACTTCGACAGCAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT 300
QY 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGCAAGAAATCTTGACCTACTATATAATAA 360
QY 361 GGTCTACATGCTAATGGGAATATGGAAATGCAAGTGCAGACAGCTCAAACTACTATGGTAA 420
Db 361 GGTCTACATGCTAATGGGAATATGGAAATGCAAGTGCAGACAGCTCAAACTACTATGGTAA 420
QY 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTGGCTGGAATGCCCTCAGGCACC 480
Db 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTGGCTGGAATGCCCTCAGGCACC 480
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCCGCGAAACTTGGTCTT 540
Db 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGACCCGCGAAACTTGGTCTT 540
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTCAACAGATGAGAAGCAGTCAATAC 600
Db 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTCAACAGATGAGAAGCAGTCAATAC 600
QY 601 ACCAATTTACTGATGAGTACAAAGTCTCAAAATCAGCAAGTATTTACCTGCTTACATGGA 660
Db 601 ACCAATTTACTGATGAGTACAAAGTCTCAAAATCAGCAAGTATTTACCTGCTTACATGGA 660
QY 661 TAAATTTACCTCAAGGAAGTCAATCAAGTTGAAGCAAGATGAGTCAAGTCACTAC 720
Db 661 TAAATTTACCTCAAGGAAGTCAATCAAGTTGAAGCAAGATGAGTCAAGTCACTAC 720
QY 721 AACTGGGATGATGCTACACAAATGTAGACCAAGCTCAAAACATCTCTGGGATAT 780
Db 721 AACTGGGATGATGCTACACAAATGTAGACCAAGCTCAAAACATCTCTGGGATAT 780
QY 781 TTACAATACAGCAAGTACGTTGCTTATCCAGACGATGAATTCGAAGTTCCTTACCAT 840
Db 781 TTACAATACAGCAAGTACGTTGCTTATCCAGACGATGAATTCGAAGTTCCTTACCAT 840
QY 841 TGTTCATCTTTACCGTAAAGTCAATTTGCCAGCTAGAGCAGCCCATCAGTCAAGTAA 900
Db 841 TGTTCATCTTTACCGTAAAGTCAATTTGCCAGCTAGAGCAGCCCATCAGTCAAGTAA 900
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTGAAGCAAAACCGGACTGGGATCAACTATGAA 960
Db 901 TGTTCCTTCGGAATTAACCAAGCAGTGAAGCAAAACCGGACTGGGATCAACTATGAA 960
QY 961 ACCGATACAGCAATATGCTTCGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
Db 961 ACCGATACAGCAATATGCTTCGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
QY 1021 CGTTCACGATGAGCCCTATACCTACCTCCGAGCAATACTCTCTGTTTATACTGGGATAG 1080
Db 1021 CGTTCACGATGAGCCCTATACCTACCTCCGAGCAATACTCTCTGTTTATACTGGGATAG 1080
QY 1081 GGGCTACTTTGGCAACATCACTCTGCAATACGCTCGCAACAAATCGCGAAAGTCCCGAGC 1140
Db 1081 GGGCTACTTTGGCAACATCACTCTGCAATACGCTCGCAACAAATCGCGAAAGTCCCGAGC 1140

QY 1141 CGTGGAACTCTAAACAAAGTCGGACTCAACCGCCCAAGACTTTCCTAAATGGTCTAGG 1200
Db 1141 CGTGGAACTCTAAACAAAGTCGGACTCAACCGCCCAAGACTTTCCTAAATGGTCTAGG 1200
QY 1201 AATCAGCTACCAAGTATTTCACTACTCAATGCCATTTCAAGTAAACACCAACGAATCAGA 1260
Db 1201 AATCAGCTACCAAGTATTTCACTACTCAATGCCATTTCAAGTAAACACCAACGAATCAGA 1260
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAGTGGCTGCTGCTTACGGTTCGCTTTGCAAAATGG 1320
Db 1261 CAAAAAATATGAGCAAGTAGTGAAGAGTGGCTGCTGCTTACGGTTCGCTTTGCAAAATGG 1320
QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGTAGGGAGTGAATA 1380
Db 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGTAGGGAGTGAATA 1380
QY 1381 AGAGTTCTCTAATGTCGGAACCTGTCGCCATGAAGAAAGCAGCAGCCTATATGATGACCGA 1440
Db 1381 AGAGTTCTCTAATGTCGGAACCTGTCGCCATGAAGAAAGCAGCAGCCTATATGATGACCGA 1440
QY 1441 CATGATGAAAACAGTCTTGACTTATGGAATCGAGCAAAATGCTATCTTGTGCTGGCTCCC 1500
Db 1441 CATGATGAAAACAGTCTTGACTTATGGAATCGAGCAAAATGCTATCTTGTGCTGGCTCCC 1500
QY 1501 TCAGGCTGGTAAACAGGAACTCTAATATACAGACGAGGAAATTTGAAACACACATCAA 1560
Db 1501 TCAGGCTGGTAAACAGGAACTCTAATATACAGACGAGGAAATTTGAAACACACATCAA 1560
QY 1561 GACCTCTCAATTTGTAGCACTGATGAATTTTGTGCTGCTATACGCTAAATATTCAT 1620
Db 1561 GACCTCTCAATTTGTAGCACTGATGAATTTTGTGCTGCTATACGCTAAATATTCAT 1620
QY 1621 GGCTGTATGGACAGGCTATTCTAAACGCTGACACACCTTTGAGCAATGGCTTACGCT 1680
Db 1621 GGCTGTATGGACAGGCTATTCTAAACGCTGACACACCTTTGAGCAATGGCTTACGCT 1680
QY 1681 CGCTGCCAAAAGTTTACCCTCTATGATGACCTACCTGCTGAAGGAAGCAATCCAGAGA 1740
Db 1681 CGCTGCCAAAAGTTTACCCTCTATGATGACCTACCTGCTGAAGGAAGCAATCCAGAGA 1740
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATTTAAATAATGGTCTCG 1800
Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATTTAAATAATGGTCTCG 1800
QY 1801 TTCTACGTGGAACTCACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
Db 1801 TTCTACGTGGAACTCACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
QY 1861 ATCAGATAGTTCAACTTCAAGTCTACAGTCTAGCTCAACACTTCCAGCAAAATATAGTACGAC 1920
Db 1861 ATCAGATAGTTCAACTTCAAGTCTAGCTCAACACTTCCAGCAAAATATAGTACGAC 1920
QY 1921 TACCAATCTTACAAATATACGCAAAATCAATCAACCCCTGATCAACAAATCAGAA 1980
Db 1921 TACCAATCTTACAAATATACGCAAAATCAATCAACCCCTGATCAACAAATCAGAA 1980
QY 1981 TCCTCAACAGCAGCAACCA 1999
Db 1981 TCCTCAACAGCAGCAACCA 1999

RESULT 2
US-08-481-435-5
; Sequence 5, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2049 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: PM 1
IMMEDIATE SOURCE:
LIBRARY: PCR cloning
CLONE: pARC 0512 Soluble PBP 1A del 38
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2049
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2046
US-08-481-435-5

Query Match	99.1%	Score 1981.4	DB 3	Length 2049	
Best Local Similarity	99.4%	Pred. No. 0			
Matches 1988	Conservative	0	Mismatches 11	Indels 0	Gaps
Qy 1	TAAATCTACGACAATAAAAAATCAACTCATTTGCTGTACTTGGGTTCTGAACGGCGCGTCAA	60			
Db 48	TAAATCTACGACAATAAAAAATCAACTCATTTGCTGTACTTGGGTTCTGAACGGCGCGTCAA	107			
Qy 61	TGCCCAAGCTTAATGATATTCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA	120			
Db 108	TGCCCAAGCTTAATGATATTCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA	167			
Qy 121	TCGCTTCTTCGACCACAGGGGGATTGATACATCCGATATCCTGGGAGCTTTCTTGGCGAA	180			
Db 168	TCGCTTCTTCGACCACAGGGGGATTGATACATCCGATATCCTGGGAGCTTTCTTGGCGAA	227			
Qy 181	TCGCAAGAAGCAATTCCTTCCAAGTGGATCAACTCTCACCCACACAGTTGATTAAAGTTGAC	240			
Db 228	TCTGCAAGAAGCAATTCCTTCCAAGTGGATCAGCTCTCACTCAACAGTGTATTAAAGTTGAC	287			
Qy 241	TTACTTTTCAACTTCGACTTCCGACAGAGCTATTTCCTTAAGGCTTCAGGAGCTTGGTT	300			
Db 288	TTACTTTTCAACTTCGACTTCCGACAGAGCTATTTCCTTAAGGCTTCAGGAGCTTGGTT	347			
Qy 301	AGCGATTTCAGTTAGACAAAAAGCAACCAAGCAAGAAAAATCTTGACCTACTATATAATAAA	360			

Db	348	AGCGATTACAGTTAGAAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	407
Qy	361	GGTCTACATCTCTAATGGGAACATATGGAATGAGACACAGCTCAAAACTACTATGTGTA	420
Db	408	GGTCTACATCTCTAATGGGAACATATGGAATGAGACACAGCTCAAAACTACTATGTGTA	467
Qy	421	AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCCTCAGGCACC	480
Db	468	AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCCTCAGGCACC	527
Qy	481	AAACCAATATGACCCCTATTTCACATCCAGAACGCCCAAGACGCCGCGAAACTTGGTCTT	540
Db	528	AAACCAATATGACCCCTATTTCACATCCAGAACGCCCAAGACGCCGCGAAACTTGGTCTT	587
Qy	541	ATCTGAAATCAAAATCAAGGCTACATCTCTGCTGAACAGATATGAGAAGCAGTCAATAC	600
Db	588	ATCTGAAATCAAAATCAAGGCTACATCTCTGCTGAACAGATATGAGAAGCAGTCAATAC	647
Qy	601	ACCAATTACTGATGGACTACAAGTCTCAAATCAGCAAGTAATTACCTGCTTACATGGA	660
Db	648	ACCAATTACTGATGGCTACAAAGTCTCAAATCAGCAAGTAATTACCTGCTTACATGGA	707
Qy	661	TAATTACTCTCAAGGAAGTCAATCAATCAAGTTTGAAGAAGAACAGGCTATACCTACTCAC	720
Db	708	TAATTACTCTCAAGGAAGTCAATCAATCAAGTTTGAAGAAGAACAGGCTATACCTACTCAC	767
Qy	721	AACCTGGGATGGTCTACACAATCTACACAAGCTACACAAGACTCAAAACATCTGTGGGATAT	780
Db	768	AACCTGGGATGGATGTCTACACAATGTACACAAGCTACACAAGCTCAAAACATCTGTGGGATAT	827
Qy	781	TTACAATACAGACGAATAGTGTGCCTATCCAGACAGTGAATTGCAAGTCGCTTCTACCAT	840
Db	828	TTACAATACAGACGAATAGTGTGCCTATCCAGACAGTGAATTGCAAGTCGCTTCTACCAT	887
Qy	841	TGTTGATGTTCTTAACGGTAAGTCAATGCCCAGCTAGAGCAGGCCATCAGTCAAGTAA	900
Db	888	TGTTGATGTTCTTAACGGTAAGTCAATGCCCAGCTAGAGCAGGCCATCAGTCAAGTAA	947
Qy	901	TGTTTCCCTCGGAATTAACCAAGCAGTCAAAACAACCGCAGCTGGGATCAACTATGAA	960
Db	948	TGTTTCCCTCGGAATTAACCAAGCAGTCAAAACAACCGCAGCTGGGATCAACTATGAA	1007
Qy	961	ACCGATCACAGACTATGCTCCTGCCTTGAGTACGGTGTACAGATTCAACTGCTACTAT	1020
Db	1008	ACCGATCACAGACTATGCTCCTGCCTTGAGTACGGTGTACAGATTCAACTGCTACTAT	1067
Qy	1021	CGTTACGATGAGCCCTATAACTACCCCTGGGACAATACTCCTGTTTATACTGGGATAG	1080
Db	1068	CGTTACGATGAGCCCTATAACTACCCCTGGGACAATACCCCTGTTTATACTGGGATAG	1127
Qy	1081	GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGCAAAACGTCGCCAGC	1140
Db	1128	GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGCAAAACGTCGCCAGC	1187
Qy	1141	CGTGGAAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCCTTAATGTGCTAGG	1200
Db	1188	CGTGGAAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCCTTAATGTGCTCGG	1247
Qy	1201	AATCGACTACCCCAAGTATTCACTACTCAAAATGCCATTTCAGTAACACACCGAATCAGA	1260
Db	1248	AATCGACTACCCCAAGTATTCACTACTCAAAATGCCATTTCAGTAACACACCGAATCAGA	1307
Qy	1261	CAAAAAATATGGAGCAAGTAGTGAAAGATGCGTCTGCTTACCGTGCCTTTGCAATAGG	1320
Db	1308	CAAAAAATATGGAGCAAGTAGTGAAAGATGCGTCTGCTTACCGTGCCTTTGCAATAGG	1367
Qy	1321	TGGAACCTTACTATAAAACCAATGTATATCATAAAGTCGCTTTTACTGATGGAGTGAAAA	1380
Db	1368	TGGAACCTTACTATAAAACCAATGTATATCATAAAGTCGCTTTTACTGATGGAGTGAAAA	1427
Qy	1381	AGAGTTCTCTAATGTGCGGAACCTCGTGCCATGAAGGAAACGACAGGCTATATGATGACCGA	1440

Db	1426	AGAGTTCTCTAATGTGCGAACTCTGGCATGAAAGAAACGACAGCCTATATGATGACCGA	1487
Qy	1441	CATGATGAAAACAGTCTTGACTTATGGAACTGGACGAAATGCCTATCTTGCTTGGCTCCC	1500
Db	1488	CATGATGAAAACAGTCTTGAGTTATGGAACCTGGACGAAATGCCTATCTTGCTTGGCTCCC	1547
Qy	1501	TCAGGCTGGTAAACAGGAACCTCTAACTATACAGACGAGGAAATTTGAAACACATCAA	1560
Db	1548	TCAGGCTGGTAAACAGGAACCTCTAACTATACAGACGAGGAAATTTGAAACACATCAA	1607
Qy	1561	GACCTCTCAATTTGTAGCACCTTGATGAACATATTTGCTGGCTATACGGCTAAATATTCAT	1620
Db	1608	GACCTCTCAATTTGTAGCACCTTGATGAACATATTTGCTGGCTATACGGCTAAATATTCAT	1667
Qy	1621	GGCTGATGGACAGGCTATTCTTAACCGTCTGACACCACTTTGTAGGCAATGSCCTTACGGT	1680
Db	1668	GGCTGATGGACAGGCTATTCTTAACCGTCTGACACCACTTTGTAGGCAATGSCCTTACGGT	1727
Qy	1681	CGCTGCCAAGTTTACCGTCTATGATGACCTACCTGTCTCAAGGAAGCAATCCAGAGA	1740
Db	1728	CGCTGCCAAGTTTACCGTCTATGATGACCTACCTGTCTCAAGGAAGCAATCCAGAGA	1787
Qy	1741	TTGGAATATACCAGAGGGGCTCTACAGAAATGGAGAATTCGTATTTAAAATGCTGCTCG	1800
Db	1788	TTGGAATATACCAGAGGGGCTCTACAGAAATGGAGAATTCGTATTTAAAATGCTGCTCG	1847
Qy	1801	TTCTAGTGGAACTCACTGTCTCCACAAACACCCCACTCAACTGAAAGTTCAAGCTCAATC	1860
Db	1848	TTCTAGTGGAGTCACTGTCTCCACAAACACCCCACTCAACTGAAAGTTCAAGCTCAATC	1907
Qy	1861	ATCAGATAGTTTCAACTTTCACGCTTAGCTCAACCACTCCAGACACAATTAATAGTAGGAC	1920
Db	1908	ATCAGATAGTTTCAACTTTCACGCTTAGCTCAACCACTCCAGACACAATTAATAGTAGGAC	1967
Qy	1921	TACCAATTCCTAACAATAATACGCAACAATCAAAATACAACCCCTGATCAACAAAATCAGAA	1980
Db	1968	TACCAATTCCTAACAATAATACGCAACAATCAAAATACAACCCCTGATCAACAAAATCAGAA	2027
Qy	1981	TCCTCAACCGACACAACCA	1999
Db	2028	TCCTCAACCGACACAACCA	2046

RESULTS

```

RESOL 3
US-08-245-511-3
; Sequence 3, Application US/08245511
; Patent No. 5928900
;
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
;
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 58
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 960 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: R6
IMMEDIATE SOURCE:
CLONE: SPRU42
FEATURE:
NAME/KEY: CDS
LOCATION: 1..960
US-08-245-511-3

Query Match 47.28; Score 944.2; DB 2; Length 960;
Best Local Similarity 99.78; Pred. No. 1.9e-278;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATAAAATCAACTCATTTGCTGACTTGGGTTCTGAACGCCGCGTCAA 60
DB 12 TAAATCTACGACAAATAAAATCAACTCATTTGCTGACTTGGGTTCTGAACGCCGCGTCAA 71
QY 61 TGCCCAAGCTAATGATATCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
DB 72 TGCCCAAGCTAATGATATCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 131
QY 121 TCGTCTTCTGACACAGGGGGGATGTACCATCGGTATCCTGGGAGCTTCTTTGCCCAA 180
DB 132 TCGTCTTCTGACACAGGGGGGATGTACCATCGGTATCCTGGGAGCTTCTTTGCCCAA 191
QY 181 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCCACAGTTGTATTAAGTTGAC 240
DB 192 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCTCAACAGTTGATTAAGTTGAC 251
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
DB 252 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 311
QY 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAGAGAATCTTGCACCTACTATATAATAA 360
DB 312 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAGAGAATCTTGCACCTACTATATAATAA 371
QY 361 GGCTTCATGCTCTAATGGGAATCTGGAATCGACAGCAGCTCAAAACACTACTATGTAA 420
DB 372 GGCTTCATGCTCTAATGGGAATCTGGAATCGACAGCAGCTCAAAACACTACTATGTAA 431
QY 421 AGAGCTCAATTAATTAAGTTTACCTTCAGTTAGCCCTTGTGGCTGGAATGCTCAGGCACC 480
DB 432 AGAGCTCAATTAATTAAGTTTACCTTCAGTTAGCCCTTGTGGCTGGAATGCTCAGGCACC 491
QY 481 AAACCAATATGACCCCCATTTCACATCCAGAGAGCCCAAGACCGCCGAAACTTTGGTCTT 540
DB 492 AAACCAATATGACCCCCATTTCACATCCAGAGAGCCCAAGACCGCCGAAACTTTGGTCTT 551
QY 541 ATCTGGAATGAAATCAAGGCTACATCTCGCTGAACAGTATCAGAAAGCAGTCAATAC 600
DB 552 ATCTGGAATGAAATCAAGGCTACATCTCGCTGAACAGTATCAGAAAGCAGTCAATAC 611
QY 601 ACCAATTACTGATGGACTTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA 660

```

Db 612 ACCAATTACTGATGGGCTACAAAGTCTCAATCAGCAAGTAATTACCCTGCTTACATGGA 671
QY 661 TAATTACCTCAAGGAAGTATCATCAATCAAGTTGAAGAGAAACAGGCTATAACTACTCTCAC 720
Db 672 TAATTACCTCAAGGAAGTATCATCAATCAAGTTGAAGAGAAACAGGCTATAACTACTCTCAC 731
QY 721 AACTGGGATGGATGCTCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
Db 732 AACTGGGATGGATGCTCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 791
QY 781 TTACAATACAGCAAGTAAGTTCCTCCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
Db 792 TTACAATACAGCAAGTAAGTTCCTCCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 851
QY 841 TGTGTATGTTTCTAAGCGTAAAGTCAATGCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 900
Db 852 TGTGTATGTTTCTAAGCGTAAAGTCAATGCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 911
QY 901 TGTTCCTTCGGGAATTAACCAAGCAGTAGAACAACACCGGACTGGGGA 949
Db 912 TGTTCCTTCGGGAATTAACCAAGCAGTAGAACAACACCGGACTGGGGA 960

RESULT 4

US-08-600-993A-3
; Sequence 3, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:

; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994

; CLASSIFICATION: 435
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994

; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 960 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6
; IMMEDIATE SOURCE:
; CLONE: SPRU42
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..960
; US-08-600-993A-3

Query Match 47.2%; Score 944.2; DB 2; Length 960;
Best Local Similarity 99.7%; Pred. No. 1.9e-278;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTAGGACAATAAAATCAACTCATGCTGACTTGGGTTCTGAAGCGCGGTCAA 60
Db 12 TAAATCTAGGACAATAAAATCAACTCATGCTGACTTGGGTTCTGAAGCGCGGTCAA 71
QY 61 TGCCCAAGCTAATGATATTCCCAACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 72 TGCCCAAGCTAATGATATTCCCAACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 131
QY 121 TCGCTTCTTCGACCACAGGGGATTTGATACCATCCGTATCTCTGGGAGCTTCTTGGCGCA 180
Db 132 TCGCTTCTTCGACCACAGGGGATTTGATACCATCCGTATCTCTGGGAGCTTCTTGGCGCA 191
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCAACCAACAGTTGATTAAGTTGAC 240
Db 192 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCAACCAACAGTTGATTAAGTTGAC 251
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 252 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 311
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGCACCTACTATATAATAA 360
Db 312 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGCACCTACTATATAATAA 371
QY 361 GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGCAGCTCAAACTACTATGGTAA 420
Db 372 GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGCAGCTCAAACTACTATGGTAA 431
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGTGGTGAATGCCTCAGGCACC 480
Db 432 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGTGGTGAATGCCTCAGGCACC 491
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCGCGGAAACTTGGTCTT 540
Db 492 AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCGCGGAAACTTGGTCTT 551
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 600
Db 552 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 611
QY 601 ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAAATTCCTGCTTACATGGA 660
Db 612 ACCAATTACTGATGGGCTACAAAGTCTCAAAATCAGCAAGTAAATTCCTGCTTACATGGA 671
QY 661 TAATTACCTCAAGGAAGTATCAATCAAGTTGAAGAGAAACAGGCTATAACCTACTCAC 720
Db 672 TAATTACCTCAAGGAAGTATCAATCAAGTTGAAGAGAAACAGGCTATAACCTACTCAC 731
QY 721 AACTGGGATGGATGCTCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
Db 732 AACTGGGATGGATGCTCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 791
QY 781 TTACAATACAGCAAGTAAGTTCCTCCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
Db 792 TTACAATACAGCAAGTAAGTTCCTCCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 851

QY 841 TGTGTGATGTTTCAACGGTAAAGTCAATCCCGAGCTAGAGCAACGCCATCAGTCAAGTAA 900
Db 852 TTTGTATGTTTCAACGGTAAAGTCAATCCCGAGCTAGAGCAACGCCATCAGTCAAGTAA 911
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACACCGGACTGGGA 949
Db 912 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACACCGGACTGGGA 960

RESULT 5
US-08-743-637B-27/C
; Sequence 27, Application US/08743637B
; Patent No. 5994066
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: PICARD, Francois J.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 EAST WISCONSIN AVENUE
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,637B
; FILING DATE: 04-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/526,840
; FILING DATE: 11-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 850586.90012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5000
; TELEFAX: (414) 277-5591
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Haemophilus influenzae
US-08-743-637B-27

Query Match 5.9%; Score 117.2; DB 2; Length 9100;
Best Local Similarity 50.3%; Pred. No. 2e-25;
Matches 349; Conservative 0; Mismatches 333; Indels 12; Gaps 2;

QY 22 TCAACTATGTCGACTGGTGTGAACGCCGCGTCAATGCCCAAGCTAAATGATATCC 81
Db 3498 TAAATTAATGGGAAGTGGGTGAGCAACGCCGTATTCCAGTGAATAGCCGATGTGCC 3439
QY 82 CACAGATTGGTAAAGCAATCGTTCTATCGAAGCACCATCGCTTCGACCAAGGG 141
Db 3438 ACAACGCTTAATGCGCATTTTACGCGAAGAACAGCAGTCGTTTTACGATCATCACGG 3379

QY 142 GATTGATACCATCCGCTATCCTCGGAGCTTTCTTTCGCCAATCTGCAAA---GCAATTCCT 198
Db 3378 ATTAGACCCCTATCGGCTATGCGCGTCAATGTTGTCGCGAGTAAATGCGGTCATC 3319
QY 199 CCAAGGTGGATCAACTCTCACCACACAGTGTATTAAGTTGACTTACTTTCACACTTCGAC 258
Db 3318 ACAAGGCGCAAGTACGATTACTCAACAAITAGGCGCTAACTTTTCTTA-----AC 3368
QY 259 TTCCGACCAGACTATTCTCTGAAGCTCAGGAAGCTTGTTAGCGATTCAAGTTAGAACA 318
Db 3267 CTCAGAAAAACCATTTATTCGTAAGCTCGTGAAGCGCTGTCGCGTAGAATCGAAA 3208
QY 319 AAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAATAAGGTCTACATGTCTAATGG 378
Db 3207 TACTCTCAACAAACAAGAAATATTAGAGCTTTATTTAAACAATACTTTTAGGCTATCG 3148
QY 379 GAACATATGAATGCAGACAGCAGCTCAAAACTACTATGTAAAGACCTCAATAATTAAG 438
Db 3147 TTCTTATGTTGTCAGCGCGACACAAACCTATTTTCGTTAAATCAATGAATGAATTGAC 3088
QY 439 TTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTCAGGACCAACCAATATGACCCCTA 498
Db 3087 CTATCGGAATGGCGATTATTGCTGTTTACCTAAGCACCTTCAACATGACCCGCT 3028
QY 499 TTCACATCCAGAACGACGCCAAGACCGCGAACTTTGGTCTTTATCTGAAATGAAATCA 558
Db 3027 TTATTTCTTTAAAACGTTTCAGAGAAGCGCGCAATGTTGCTTAAGCCGTATGTTAGATGA 2968
QY 559 AGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATACACCACTACTATGACT 618
Db 2967 AAAATACATCAGCAAAAGAATAATGCTGCATTGAAAGAGCCGATTGTGCGGAGCTA 2908
QY 619 ACAAGTCTCAAAATCAGCAAGTAAATTAACCTGCTTACATGGATAATTAACCTCAAGGAAT 678
Db 2907 TCAGCGCGCAAAATTTGAATTTTCGAGCCGATTTATGCTCACTGAATGCTGCGTCAAGAAAT 2848
QY 679 CATCAATCAAGTTGAAGAAGAACACAGGCTATAAC 712
Db 2847 GGTGCTGCTGTTTGGCGAAGAAAAATGCTTACACC 2814

RESULT 6
US-08-526-840B-27/C
; Sequence 27, Application US/08526840B
; Patent No. 6001564
; GENERAL INFORMATION:
; APPLICANT: BERGERON, Michel G.
; APPLICANT: OUELLETTE, Marc
; APPLICANT: ROY, Paul H.
; TITLE OF INVENTION: SPECIFIC AND UNIVERSAL PROBES AND
; TITLE OF INVENTION: AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY
; TITLE OF INVENTION: COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE
; TITLE OF INVENTION: FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...
; NUMBER OF SEQUENCES: 177
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: QUARLES & BRADY
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/526,840B
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/304,732
; FILING DATE: 12-SEP-1994

Db 200 GCTTCTACGAGCATACGCGGTGGGATCTCCGTCGAGCAAGCGTGGCGC 259
QY 183 TGCAGAGCAATCCCTCCCAAGGTGGATCAACTCTACCCCAACAGATTGATTAAGTTGACTT 242
Db 260 TGTCTCCGGTCACGGCTCAAGGGGCAAGTACCATTACCAGCAGCTGGCGAGAACT 319
QY 243 ACTTTTCAACTTCGACTTCGACCAAGACTATTCTCGTAAGGCTCAGGAAGCTTGGTTAG 302
Db 320 TCTTCTCAGTCC-----AGAACGCACGCTGATGCGTAAGATTAAAGGAAGTCTTCTCTCG 373
QY 303 CGATTTCAGTTAGCAACAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAAGG 362
Db 374 CGATTTCGATTAAGCACTGCTGACGAAAGACGAGATCTCGAGCTTTATCTGAACAAGA 433
QY 363 TCTACATGTCTAATGGAACTATGGAATGCAGACAGCAGCTCAAAACTACTATGTGTAAG 422
Db 434 TTTACCTTGGTTACCGCGCTATGGTGTGCTGCGGCAACAAGTCTATTTTCGGAAAAA 493
QY 423 ACCTCAATTAATTAAGTTTACCTCAGTTAGCCTTGTGCTGGAATGSCCTCAGGCACCAA 482
Db 494 CGGTCCACCAACTGACGCTGAACGAATGCGGTGATAGCGGCTGCCGAAAGCGCTT 553
QY 483 ACCAATATACCCCTATTTCATCTCAGAACAGCAGCCCAACAGCCGCAAACTTGGTCTTAT 542
Db 554 CCACCTTCAACCCGCTCTACTCGATGGATCGTCCGCTCGCGGCGTAAACGTCGTCTGT 613
QY 543 CTGAATGAAATCAATCAAGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATACAC 602
Db 614 CGCGGATGCTGATGAAGGTTATATACCCCAACAGCTTCGATCAGACAGCAGCTGAGG 673
QY 603 CAATTAAGT 613
Db 674 CGATTAACGCT 684

RESULT 8
US-08-771-716-1
; Sequence 1, Application US/08771716
; Patent No. 5922540
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S.Richard
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Monofunctional Glycosyltransferase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,716
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..807
US-08-771-716-1

Query Match 3.1%; Score 62.2; DB 2; Length 807;
Best Local Similarity 49.9%; Pred. No. 3.4e-09;
Matches 191; Conservative 0; Mismatches 183; Indels 9; Gaps 1;

QY 67 AGCTATGATATCCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACCATCGCTT 126
Db 252 AGCTGATAACATGCCAGAGTATGTTAAAGTGCTTTATTTCAATGAAGATGAACGATT 311
QY 127 CTTGACACAGGGGATGATACCATCGTATCTCTGGGAGCTTCTTGGCAATCTGCA 186
Db 312 CTACATCATCATGATTGATTTGAAGGTACAACTAGAGCTTTATTTCAAGATTAG 371
QY 187 AAGCAATTCCTCCCAAGGTGGATCAACTCTCACCACAGTTGATTAAGTTGACTTACTT 246
Db 372 CGACAGAGATGTCAAGGTGGTAGTACCATTACACAAAGTTGTCAAAAATTTATTTTA 431
QY 247 TTCACTTCGACTTCGACAGGACTATTTCTCGTAGGCTCAGGAAGCTTGGTTAGCAT 306
Db 432 TGATAAT-----GATCGTTTACTTACTAGAAAGTAAAGAATTAATTTTAGTCA 482
QY 307 TCAGTTAGAAACAAAAGCAACCAAGAAATCTTGACCTACTATATAAAGGTCTA 366
Db 483 TCGAGTTGAAAACAATATAATAAGACGAAATTTAAGCTTTTATTAATAATATTTA 542
QY 367 CATGCTTAATGGAACTATGGAATGCAGACAGCAGCTCAAAACTACTATGTTAAAGACCT 426
Db 543 CTTTGGGGATAATCAATATACGCTTCGAGGCGCAGCAAAACCATTTCTTGGAAACACGT 602
QY 427 CAATAATTTAAGTTTACCTCAGT 449
Db 603 GAATAAAAATAGTACAAACATGT 625

RESULT 9
US-08-771-716-3
; Sequence 3, Application US/08771716
; Patent No. 5922540
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S.Richard
; APPLICANT: Skatrud, Paul L.
; TITLE OF INVENTION: Monofunctional Glycosyltransferase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/771,716
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 807 base pairs
; TYPE: nucleic acid

STREET: Lilly Corporate Center
 CITY: Indianapolis
 STATE: Indiana
 COUNTRY: U.S.
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/057,720A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Webster, Thomas D.
 REGISTRATION NUMBER: 39,872
 REFERENCE/DOCKET NUMBER: X-11067
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-276-3334
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 807 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-09-057-720A-3

Query Match 3.1%; Score 62.2; DB 3; Length 807;
 Best Local Similarity 33.9%; Pred. No. 3.4e-09;
 Matches 130; Conservative 61; Mismatches 183; Indels 9; Gaps 1;
 QY 67 AGCTATGATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCATCGCTT 126
 DB 252 ACCUGAUAACUGCCAGAGUAUGUUAAGGUGCCUUUAUUAUGAAGUAGCAUAU 311
 QY 127 CTTGACACAGGGGATTGATACATCCGTATCTCTGGGAGCTTCTTCGCGCAATCTGA 186
 DB 312 CUACAUAUCAUGGAGUUGCUUUAAGAGGUAACUAGAGGCUUUAUUAUUAUUAU 371
 QY 187 AAGCAATCCCTCCAGGATGATCACTCTACCCCAAGTGTGATTGATTGACTTACTT 246
 DB 372 GCACAGAGUUGCAGAGGUGUAGUACCAUUAACACACAGGUGUUGUUAUUAUUAU 431
 QY 247 TTCAACTTCGACTTCGACGAGACTATTTCTGTAAGGCTCAGGAGCTTGGTTAGCGAT 306
 DB 432 UGAUAU-----GAUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 482
 QY 307 TCAGTTAGAAACAAAGCAACCAAGCAAGAAATCTTGACTACTATATAAATAAGTCTA 366
 DB 483 UCGAGUUGAAACAAUUAUUAAGAACGAAAUUAUUAAGUUUAUUAUUAUUAUUAU 542
 QY 367 CATGCTATGGAATGGAATGCAGACAGCAGCTCAAACTACTATGTTAAAGACCT 426
 DB 543 CUUUGGGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 602
 QY 427 CAATAATTTAAGTTTACCTCAGT 449
 DB 603 GAAUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 625

RESULT 12
 US-08-731-716-1
 ; Sequence 1, Application US/08731716
 ; Patent No. 5789202
 ; GENERAL INFORMATION:
 ; APPLICANT: Hoskins, John
 ; APPLICANT: Jaskunas, S. Richard
 ; APPLICANT: Rockey, Pamela K.

APPLICANT: Zhao, Genshi
 APPLICANT: Rostock, Paul R. Jr.
 APPLICANT: No. 5789202is, Franklin H.
 TITLE OF INVENTION: Penicillin Binding Protein From
 TITLE OF INVENTION: Streptococcus Pneumoniae
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Eli Lilly and Company
 STREET: Lilly Corporate Center
 CITY: Indianapolis
 STATE: Indiana
 COUNTRY: U.S.
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/731,716
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Webster, Thomas D.
 REGISTRATION NUMBER: 39,872
 REFERENCE/DOCKET NUMBER: X-10,887
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-276-3334
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2193 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..2193
 US-08-731-716-1

Query Match 2.8%; Score 56.4; DB 1; Length 2193;
 Best Local Similarity 44.6%; Pred. No. 3.4e-07;
 Matches 381; Conservative 0; Mismatches 446; Indels 27; Gaps 3;
 QY 5 ATCTACGACAATAAAATCAACTCATTTGCTGACTTGGTTCGAGCGCCGCTCAATGCC 64
 DB 283 ATTTTGGCCGTGAAGAAAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 342
 QY 65 CAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCATCGC 124
 DB 343 GAGCTGACTGACATGATTAATAAATTCGAGATGCTGTTATTTCGACAGAACCCCTTCT 402
 QY 125 TTCTTCGACACAGGGGATGATACCATTCGCTATCCTGGAGCTTCTTTCGCAATCTG 184
 DB 403 TTCTATAAATTCAGGGGATTAACATGCGCGTTTC-----TTCTTGGCTATTGTC 453
 QY 185 CAAGCAATTCCTCCAGGTGGATCACTCTCACCAACAGTTGATTAAAGTTGACTTAC 244
 DB 454 ACTGCTGGAGCTTCAAGTGGTGGTCTACCATTAACCAACAGCTGCTAATAACGCTAT 513
 QY 245 TTTTCAACTTCGACTTCCGACAGACTATTTCGTAAGGCTCAGGAACCTTGGTTAGCG 304
 DB 514 TTAT-----CGCAGGATCAAACTGTTGAGAGAAAAGCGAAAGAAATTTTTCCTTGC 564
 QY 305 ATTGAGTTAGAAACAAACCAACCAAGCAAGAAATCTTCACCTACTATATAATAAGGTC 364
 DB 565 TTAGATTAAGCAAAAATAATAGTAGGAGCAAAATTCATACCATGATACCTTAACACGCT 624
 QY 365 TACATGCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAACTACTATGTAAGAC 424

Db 625 TATTTTGGAAATGGTGTGGGGGTAGAGATGCGAGTAAAGAAATACATTGGAGTTCT 684
QY 425 CTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGATGCTCCTCAGGCACCAAAAC 484
Db 685 GCATCAGAGTGGTCTGATCAAGCTGCACTCTGGCAGGATGCTCAAGGGCGGGA 744
QY 485 CAATATGACCCCTATTTCATCCAGAACGAGCCGACAGCCGCGAAACTTGGTCTTATCT 544
Db 745 CTGTATAATCCCTTGAATTCCTAGAGATTTCTACTAATCGGCGGATACTGCTTGCAG 804
QY 545 GAAATGAAAAATCAGGCTACATCTCTGCTGACAGATGAGAAAGCAGTCAATACCA 604
Db 805 ATATGTTGGACGAGATATATTGATAAAACCCAGAAACCGAAGCTCTGGAAGTTGAT 864
QY 605 ATTACTGATGGACTCAAAAGTCTCAAAATCAGCAAGTAAATACCCCTCTTACATGGATAAT 664
Db 865 ATGACTTCGCAATTGCACGATAAGTATGAAGGAAAAATCTCAGATTACCGTTACCCCTCT 924
QY 665 TACCTCAAGGAAGTCAATCAAGTTG-----AAGAAGAAACAGGCTATACCTA 715
Db 925 TATTTTGAATGCGGTGGTTAATGAAGCTGTTTCCAAATATAATCTACAGAGGAAGAT 984
QY 716 CTCACAACCTGGGATGATGCTCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGTGG 775
Db 985 GTCAATAATGGCTACCGCATTTACACAGAGCTGGACCAAAACTACCAACAAATATGCA 1044
QY 776 GATATTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATGCAAGTCGCTTCT 835
Db 1045 ATTGTTTATGAACAACATCGCTATTTCGAGGCGCAGAGGATGGAACGTTGCTCAATCA 1104
QY 836 ACCATTGTTGATGT 849
Db 1105 GGAAGTGTAGCTCT 1118

RESULT 13

US-08-731-716-3
; Sequence 3, Application US/08731716
; Patent No. 5789202
; GENERAL INFORMATION:
; APPLICANT: Hoskins, JoAnn
; APPLICANT: Jaskunas, S. Richard
; APPLICANT: Rockey, Pamela K.
; APPLICANT: Zhao, Genshi
; APPLICANT: Rostek, Paul R. Jr.
; APPLICANT: No. 5789202ris, Franklin H.
; TITLE OF INVENTION: Penicillin Binding Protein From
; TITLE OF INVENTION: Streptococcus Pneumoniae
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,716
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-10,887
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 2193 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-731-716-3

Query Match 2.8%; Score 56.4; DB 1; Length 2193;
Best Local Similarity 31.6%; Pred. No. 3.4e-07;
Matches 270; Conservative 111; Mismatches 446; Indels 27; Gaps 3;

QY 5 ATCTAGCAATAAATAATCAATCTGCTGACTTGGTCTTGGTTCTGAACCGCGCTCAATGCC 64
Db 283 AUUUUUGACCGGAGAGAAAAGAGCGUGGCUUGUCUGUCAAAGGAAACCUAUGUU 342
QY 65 CAAGCTAATGATATTCCCAACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCAATCG 124
Db 343 GAGCUGACUGACAUCAGUAAAAACUUGCAGAAUGUGUUAUUGCGACAGAAACCGUUCU 402
QY 125 TTCTTCGACCACAGGGGGGATTGATACCATCCGTTATCTCTGGAGCTTTCTTGGCAATCTG 184
Db 403 UUCUUAUAAAAUAGACGGGAUUAACUUAUGCCGUUUC-----UUCUUGGCUAUGUC 453
QY 185 CAAAGCAATTCCTCCCAAGTGGATCAACTCTCACCAACAGTGTGATTAAAGTTGACTTAC 244
Db 454 ACUGUGGAGGCUUACAGUGUGGUCUACCAUUAUCCCAACAGCUGGCUAAAAACGCCU 513
QY 245 TTTTCAACTTCGACTTCCAGCAGACTATTCTCTGTAAGGCTCAGGAAGCTTGGTTAGCG 304
Db 514 UUAU-----CGCAGGAUCAACUGUUGAGAGAAAAAGCGAAAGAAUUUUUCCUUGCC 564
QY 305 ATTCAGTTAGAACAAAAACCAACAGCAAGAAATCTTCACCTACTATATAATAAGTGC 364
Db 565 UUGAGAAUUAAGCAAAAAAUUAUAGAGGAGCAAAUUAUACCAUUGUACCUAACAACGCU 624
QY 365 TACATGCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAAACTACTATGTAAGAC 424
Db 625 UAUUUUGAAUUGGUGUGUGGUGUAGAAGUAGUAGAAUAUACUUGGAGUUUCU 684
QY 425 CTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGTGAATGCTCCTCAGGCACCAAC 484
Db 685 GCAUCAGAGAUGAGUGUGGAUCAAAGCUGGACUCUGGCAGGGAUGUCAAGGGCGGAA 744
QY 485 CAATATGACCCCTATTTCATCCAGAACGAGCCCAAGACCGCGGAACTTGGTCTTATCT 544
Db 745 CUGUAUAUCCCUUGAAUUCUGUAGAAUUAUUAUUCGCGCGGAUACUGUCUUGCGAG 804
QY 545 GAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATACCA 604
Db 805 AAUUGGUUGCAGCAGGAUAUAUUAUAAACCAAGAAACCGAGCUCUGAAGUUGAU 864
QY 605 ATTACTGATGGACTCAAAAGTCTCAAAATCAGCAAGTAAATTCCTGCTTACATGGATAAT 664
Db 865 AUGACUUCGCAAUUGCAGGAUUAUGAAGGAAAAAUCAGAUUACCUUACCCUUCU 924
QY 665 TACCTCAAGGAAGTCAATCAAGTTG-----AAGAAGAAACAGGCTATACCTA 715
Db 925 UAUUUUGCGGUGUUAUUAUGAAGCUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 984
QY 716 CTCACAACCTGGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGG 775
Db 985 GUCAUUAUUGGCUUACCGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1044
QY 776 GATATTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATGCAAGTCGCTTCT 835
Db 1045 AUUGUUUAUAGAAACACACUUGCUUAUUCGAGGCGCAGAGGAAACCUUUAUUAUUA 1104
QY 836 ACCATTGTTGATGT 849
Db 1105 GGAAGUGUAGUCU 1118

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 12:37:03 ; Search time 3441.19 seconds
(without alignments)
12567.793 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTAGCACAATAAAA.....ATCCTCAACGACACAACCA 1999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 21979536 seqs, 10817449327 residues
Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:*

2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*

3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*

4: /cgn2_6/ptodata/2/pna/US08_COMB.seq:*

5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*

6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*

7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*

8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*

9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*

10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*

11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*

12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*

13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*

14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*

15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*

16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*

17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*

18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*

19: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*

20: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*

21: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*

22: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*

23: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*

24: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*

25: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*

26: /cgn2_6/ptodata/2/pna/US096E_COMB.seq:*

27: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*

28: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*

29: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*

30: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*

31: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*

32: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*

33: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*

34: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*

35: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*

36: /cgn2_6/ptodata/2/pna/US100_COMB.seq:*

37: /cgn2_6/ptodata/2/pna/US101_COMB.seq:*

38: /cgn2_6/ptodata/2/pna/US6000_COMB.seq:*

39: /cgn2_6/ptodata/2/pna/US6001_COMB.seq:*

40: /cgn2_6/ptodata/2/pna/US6002_COMB.seq:*

41: /cgn2_6/ptodata/2/pna/US6003_COMB.seq:*

42: /cgn2_6/ptodata/2/pna/US6004_COMB.seq:*

43: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1999	100.0	1999	20	US-09-536-784-1	Sequence 1, Appli
2	1999	100.0	1999	30	US-09-765-271-1	Sequence 1, Appli
3	1999	100.0	1999	30	US-09-765-272-1	Sequence 1, Appli
4	1999	100.0	10383	41	US-60-029-960-81	Sequence 81, Appl
5	1999	100.0	10711	13	US-08-961-527-145	Sequence 145, App
6	1987.8	99.4	2160	22	US-09-583-110-1312	Sequence 1102, Ap
7	1987.8	99.4	2166	15	US-09-107-433-1102	Sequence 9325, Ap
8	1981.4	99.1	2160	1	PCT-US02-03987-9325	Sequence 9325, Ap
9	1981.4	99.1	2160	31	US-09-815-242-9325	Sequence 9325, Ap
10	1981.4	99.1	2160	37	US-10-072-851-9325	Sequence 596, App
11	1981.4	99.1	10333	45	US-60-061-998-596	Sequence 596, App
12	1631.2	81.6	9845	45	US-60-068-175-596	Sequence 139, App
13	944.2	47.2	960	5	US-08-116-541-3	Sequence 139, App
14	808.6	40.5	2172	1	PCT-US97-14436-139	Sequence 139, App
15	808.6	40.5	2172	13	US-08-911-503-139	Sequence 139, App
16	808.6	40.5	2172	13	US-08-911-503A-139	Sequence 6631, Ap
17	574.6	28.7	2337	1	PCT-US02-03987-6631	Sequence 6631, Ap
18	574.6	28.7	2337	31	US-09-815-242-6631	Sequence 1534, Ap
19	574.6	28.7	2337	37	US-10-072-851-6631	Sequence 217, App
20	573	28.7	2370	15	US-09-134-000-1534	Sequence 217, App
21	559	28.0	8395	14	US-09-070-927-217	Sequence 217, App
22	559	28.0	8395	14	US-09-070-927A-217	Sequence 2013, Ap
23	553	27.7	2472	15	US-09-107-532A-2013	Sequence 2013, Ap
24	553	27.7	2472	15	US-09-107-532A-2013	Sequence 495, App
25	496.2	24.8	1386	37	US-10-091-007-91	Sequence 495, App
26	410	20.5	2470	45	US-60-068-186-495	Sequence 495, App
27	410	20.5	2474	44	US-60-050-444-495	Sequence 710, App
28	352.2	17.6	3478	45	US-60-068-217-710	Sequence 14, Appl
29	348.6	17.4	3279	24	US-09-634-238-44	Sequence 108, App
30	348.6	17.4	7818	24	US-09-634-238-14	
31	273	13.7	933	43	US-60-045-649-108	

32 273 13.7 933 43 US-60-046-653-183 Sequence 183, App
33 258 12.9 323 41 US-60-029-960-934 Sequence 934, App
34 221.4 11.1 6394 26 US-09-663-779-1130 Sequence 1130, Ap
35 204 10.2 204 22 US-09-583-110-1313 Sequence 1313, Ap
36 164.2 8.2 3748 26 US-09-663-779-1038 Sequence 1038, Ap
37 163 8.2 2229 36 US-09-974-300-1671 Sequence 1671, Ap
38 123.8 6.2 1899 24 US-09-620-608-1543 Sequence 1543, Ap
39 123.8 6.2 1899 53 US-60-144-883-1543 Sequence 1543, Ap
40 123.8 6.2 2568 17 US-09-328-352-2090 Sequence 2090, Ap
41 122.8 6.1 1590 42 US-60-038-697-335 Sequence 335, App
42 120.6 6.1 1590 43 US-60-046-714-308 Sequence 308, App
43 120.6 6.0 2181 1 PCT-US02-03987-4727 Sequence 4727, Ap
44 120.6 6.0 2181 31 US-09-815-242-4727 Sequence 4727, Ap
45 120.6 6.0 2181 37 US-10-072-851-4727 Sequence 4727, Ap

ALIGNMENTS

RESULT 1

US-09-536-784-1

; Sequence 1, Application US/09536784

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESS: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/536,784

; FILING DATE: 30-Oct-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/961,083

; FILING DATE: OCT-30-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

; REFERENCE/DOCKET NUMBER: PB340P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1999 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-536-784-1

Query Match 100.0%; Score 1999; DB 20; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATAAATAAATACTATGCTGACTTGGGTTCTGAAACGGCGGTCAA 60

Db 1 TAAATCTAGCAATAAATAAATACTATGCTGACTTGGGTTCTGAAACGGCGGTCAA 60

QY 61 TCCCCAAGCTAATGATATCCACAGATTGGTTAGGCAATCGTTTCTATCGAAGACCA 120

Db 61 TCCCCAAGCTAATGATATCCACAGATTGGTTAGGCAATCGTTTCTATCGAAGACCA 120

QY 121 TCGCTTCTTCGACCAACAGGGGGATTGATACCATCCGTATCCTCTGGGAGCTTCTTGGCAA 180
Db 121 TCGCTTCTTCGACCAACAGGGGGATTGATACCATCCGTATCCTCTGGGAGCTTCTTGGCAA 180
QY 181 TCTGCAAAACAAATTCCTCCCAAGGTGGATCACTCTCACCCACACAGTTGATTAAAGTTGAC 240
Db 181 TCTGCAAAACAAATTCCTCCCAAGGTGGATCACTCTCACCCACACAGTTGATTAAAGTTGAC 240
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCGTAAGGCTCAGGAAGCTTGGTT 300
QY 301 AGCGATTTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db 301 AGCGATTTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
QY 361 GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACACTACTATGATA 420
Db 361 GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACACTACTATGATA 420
QY 421 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGATGCTCAGGCACC 480
Db 421 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGATGCTCAGGCACC 480
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACGCCGGAACACTTGGTCT 540
Db 481 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACGCCGGAACACTTGGTCT 540
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAATTAACCTGCTTACATGGA 660
Db 601 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAATTAACCTGCTTACATGGA 660
QY 661 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAAGCAAGGCTATAACCTACTCAC 720
Db 661 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAAGCAAGGCTATAACCTACTCAC 720
QY 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGGCTCAAAACACTCTGTGGGATAT 780
Db 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGGCTCAAAACACTCTGTGGGATAT 780
QY 781 TTACAATACAGAGCAATACGTTGCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
Db 781 TTACAATACAGAGCAATACGTTGCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
QY 841 TGTGATGTTTCTAACGGTAAAGTCAATGCCAGCTAGGACGACGCCATCAGTCAAGTAA 900
Db 841 TGTGATGTTTCTAACGGTAAAGTCAATGCCAGCTAGGACGACGCCATCAGTCAAGTAA 900
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCGGAGCTGGGGATCAACTATGAA 960
Db 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCGGAGCTGGGGATCAACTATGAA 960
QY 961 ACCGATCACAGACTATGCTCCTGCTGGAGTACGGTGTCTAGGATTAACCTACTACTAT 1020
Db 961 ACCGATCACAGACTATGCTCCTGCTGGAGTACGGTGTCTAGGATTAACCTACTACTAT 1020
QY 1021 CGTTACAGATGATGCCCTATAAATACCTGGGACAAATACTCCTGTTTATAAATCGGATAG 1080
Db 1021 CGTTACAGATGATGCCCTATAAATACCTGGGACAAATACTCCTGTTTATAAATCGGATAG 1080
QY 1081 GGGCTACTTGGCAACATCACCTTGGCAATACGCCCTGCAACAATCGGAAACGCTCCAGC 1140
Db 1081 GGGCTACTTGGCAACATCACCTTGGCAATACGCCCTGCAACAATCGGAAACGCTCCAGC 1140
QY 1141 CGTGGAAACTCTAAACAAGTTCGAGCTCAACCGCGCAAGACTTTTCTAAATGGTCTAGG 1200
Db 1141 CGTGGAAACTCTAAACAAGTTCGAGCTCAACCGCGCGCAAGACTTTTCTAAATGGTCTAGG 1200

QY 601 ACCAATTACTGATGGAGTACAAAGTCTCAAAATCAGCAAGTAATATACCTGCTTACATGA 660
 DB 601 ACCAATTACTGATGGAGTACAAAGTCTCAAAATCAGCAAGTAATATACCTGCTTACATGA 660
 QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAACAGGCTATAACCTACTCAC 720
 DB 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAACAGGCTATAACCTACTCAC 720
 QY 721 AACTGGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
 DB 721 AACTGGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
 QY 781 TTACAATACAGAGCAATAGCTTGCCTATCCAGACGATGAATTTGCAAGTGCCTTCTACCAT 840
 DB 781 TTACAATACAGAGCAATAGCTTGCCTATCCAGACGATGAATTTGCAAGTGCCTTCTACCAT 840
 QY 841 TGTGTGATGTTTCAACGGTAAGTCAATTTGCCAGCTAGGACGCGCATCAGTCAAGTAA 900
 DB 841 TGTGTGATGTTTCAACGGTAAGTCAATTTGCCAGCTAGGACGCGCATCAGTCAAGTAA 900
 QY 901 TGTTCCTTCGGAATTAACCAAGCAGTGAACCAACACCGGACTGGGGATCAACTATGAA 960
 DB 901 TGTTCCTTCGGAATTAACCAAGCAGTGAACCAACACCGGACTGGGGATCAACTATGAA 960
 QY 961 ACCGATACAGACTAGCTGCTGCTTGGAGTACGGTGTCTAGCATTCAACTGCTACTAT 1020
 DB 961 ACCGATACAGACTAGCTGCTGCTTGGAGTACGGTGTCTAGCATTCAACTGCTACTAT 1020
 QY 1021 CGTTCAGATGAGCCCTATAACTACCTGGGCAAAATCTCCTGTTTATAACTGGGATAG 1080
 DB 1021 CGTTCAGATGAGCCCTATAACTACCTGGGCAAAATCTCCTGTTTATAACTGGGATAG 1080
 QY 1081 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAAATCGGAAACCTCCAGC 1140
 DB 1081 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAAATCGGAAACCTCCAGC 1140
 QY 1141 CGTGGAACTCTAAACAAGTGGAGTCAACCGCGCAAGACTTTCCTAAATGGCTTAGG 1200
 DB 1141 CGTGGAACTCTAAACAAGTGGAGTCAACCGCGCAAGACTTTCCTAAATGGCTTAGG 1200
 QY 1201 AATCGACTCCCAAGTATTCACCTACTCAATGCGCTTCAAGTCAACACACGCAATCAGA 1260
 DB 1201 AATCGACTCCCAAGTATTCACCTACTCAATGCGCTTCAAGTCAACACACGCAATCAGA 1260
 QY 1261 CAAAAAATATGGAGCAAGTAGTGAAGATGGCTGCTTACGCTGCTTTGCAAAATGG 1320
 DB 1261 CAAAAAATATGGAGCAAGTAGTGAAGATGGCTGCTTACGCTGCTTTGCAAAATGG 1320
 QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGGAGTGAATA 1380
 DB 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGGAGTGAATA 1380
 QY 1381 AGAGTTCTCTAATGTCGGAACCTGCTGCCATGAAGAAACGACAGCTTATATGATACCGA 1440
 DB 1381 AGAGTTCTCTAATGTCGGAACCTGCTGCCATGAAGAAACGACAGCTTATATGATACCGA 1440
 QY 1441 CATGATGAAACAGCTTGTGACTTATGAACTGGACGAAATGCCATCTTCTTGGCTCCC 1500
 DB 1441 CATGATGAAACAGCTTGTGACTTATGAACTGGACGAAATGCCATCTTCTTGGCTCCC 1500
 QY 1501 TCAGGCTGTAAACAGGAACTCTCACTATACAGACGAGGAATTTGAAACACATCAA 1560
 DB 1501 TCAGGCTGTAAACAGGAACTCTCACTATACAGACGAGGAATTTGAAACACATCAA 1560
 QY 1561 GACCTCTCAATTTGTAGCACTGTGAATATTTGCTGCTATACGCTAAATATTTCAAT 1620
 DB 1561 GACCTCTCAATTTGTAGCACTGTGAATATTTGCTGCTATACGCTAAATATTTCAAT 1620
 QY 1621 GCGTGTATGACAGGCTATTTAAACCGCTTGACACCACTTTAGGCAATGGCCTTACGCT 1680
 DB 1621 GCGTGTATGACAGGCTATTTAAACCGCTTGACACCACTTTAGGCAATGGCCTTACGCT 1680
 QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGAAGGAAGCAATCCAGAAGA 1740

DB 1681 CGCTGCAAAAGTTTACGCTCTATGATGACCTACCTGCTGTGAAGGAAGCAATCCAGAAGA 1740
 QY 1741 TTGGAATATACAGAGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTCTCG 1800
 DB 1741 TTGGAATATACAGAGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTCTCG 1800
 QY 1801 TTCTAGCTGGAACCTACCTGCTCCACAACACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 DB 1801 TTCTAGCTGGAACCTACCTGCTCCACAACACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 QY 1861 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACACTTCCAAAGCACAATAATAGTAGCAG 1920
 DB 1861 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACACTTCCAAAGCACAATAATAGTAGCAG 1920
 QY 1921 TACCAATCTTACAATAATACGCAACAAATCAATCAACCCCTGATCAACAAATCAGAA 1980
 DB 1921 TACCAATCTTACAATAATACGCAACAAATCAATCAACCCCTGATCAACAAATCAGAA 1980
 QY 1981 TCCTCAACAGCAGCAACCA 1999
 DB 1981 TCCTCAACAGCAGCAACCA 1999
 RESULT 3
 US-09-765-272-1
 ; Sequence 1, Application US/09765272
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi et. al.
 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
 ; NUMBER OF SEQUENCES: 452
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/765,272
 ; FILING DATE: 22-Jan-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/961,083
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB340P2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1999 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-765-272-1

Query Match 100.0%; Score 1999; DB 30; Length 1999;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TAAATCTAGCAATATAAATCAACTCATGCTGACTGGTCTGCAAGCGCGCTCAA 60

Db 1 TAAATCTACGCAATAAATAAATACTCATCTGCTGACTTGGGTCTGTAACGCGCGTCAA 60
 QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 Db 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 QY 121 TCGCTTCTCGACACAGGGGATTGATACCAATCGATATCCCTGGGAGCTTCTTGGCGCAA 180
 Db 121 TCGCTTCTCGACACAGGGGATTGATACCAATCGATATCCCTGGGAGCTTCTTGGCGCAA 180
 QY 181 TCTGCAAAAGCAATTCCTCCAAAGGTGGATCAACTCTCACCACACAGTTGATTAAGTTGAC 240
 Db 181 TCTGCAAAAGCAATTCCTCCAAAGGTGGATCAACTCTCACCACACAGTTGATTAAGTTGAC 240
 QY 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
 Db 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
 QY 301 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAATAA 360
 Db 301 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAATAA 360
 QY 361 GGTCTACATGCTAATGGGAATATGGAAATGCAAGCAGCAGCTCAAACTACTATGGTAA 420
 Db 361 GGTCTACATGCTAATGGGAATATGGAAATGCAAGCAGCAGCTCAAACTACTATGGTAA 420
 QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGAAATGCCCTCAGGCACC 480
 Db 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGAAATGCCCTCAGGCACC 480
 QY 481 AAACCAATATGACCCCTATTACATCCAGAACGCCCAAGACCCGCGAAATCTGGTCTT 540
 Db 481 AAACCAATATGACCCCTATTACATCCAGAACGCCCAAGACCCGCGAAATCTGGTCTT 540
 QY 541 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
 Db 541 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
 QY 601 ACCAATTAAGTGGACTACAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGA 660
 Db 601 ACCAATTAAGTGGACTACAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGA 660
 QY 661 TAATTAACCTCAAGGAGTCAATCAAGTTGAAGAAAGCAGGCTATAACCTACTAC 720
 Db 661 TAATTAACCTCAAGGAGTCAATCAAGTTGAAGAAAGCAGGCTATAACCTACTAC 720
 QY 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT 780
 Db 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT 780
 QY 781 TTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATTTGCAAGTCCCTTACCAT 840
 Db 781 TTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATTTGCAAGTCCCTTACCAT 840
 QY 841 TGTGATGTTTCAACGTTAAAGTCAATTTGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
 Db 841 TGTGATGTTTCAACGTTAAAGTCAATTTGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
 QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGATCAACTATGAA 960
 Db 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGATCAACTATGAA 960
 QY 961 ACCGATACAGACTATGCTTCCTGCTTGGAGTACGCTGCTACGATTCAACTGCTACTAT 1020
 Db 961 ACCGATACAGACTATGCTTCCTGCTTGGAGTACGCTGCTACGATTCAACTGCTACTAT 1020
 QY 1021 CGTTCAGATGACCCCTATACTACCTGGGACAAATACTCTGTTTATACTGGGATAG 1080
 Db 1021 CGTTCAGATGACCCCTATACTACCTGGGACAAATACTCTGTTTATACTGGGATAG 1080
 QY 1081 GGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCACCAATCGGAAAGCTCCAGC 1140
 Db 1081 GGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCACCAATCGGAAAGCTCCAGC 1140

RESULT 4

US-60-029-960-81/c

; Sequence 81, Application US/60029960
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 1649
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville

QY 1141 CGTGGAACTCTAAACAAGTGGGACTCAACGCGCCAGAGACTTTCTTAAATGGTCTAGG 1200
 Db 1141 CGTGGAACTCTAAACAAGTGGGACTCAACGCGCCAGAGACTTTCTTAAATGGTCTAGG 1200
 QY 1201 AATCGACTACCAAGTATTCTACTACTCAAAATGCGATTTCAAGTAAACAACCAAGTAAATCAGA 1260
 Db 1201 AATCGACTACCAAGTATTCTACTACTCAAAATGCGATTTCAAGTAAACAACCAAGTAAATCAGA 1260
 QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAAGTGGCTGCTTACGCTGCTTTAGCGTCCCTTTGCAAAATG 1320
 Db 1261 CAAAAAATATGAGCAAGTAGTGAAGAAGTGGCTGCTTACGCTGCTTTAGCGTCCCTTTGCAAAATG 1320
 QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTAGTGATGGGAGTGAATAA 1380
 Db 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTAGTGATGGGAGTGAATAA 1380
 QY 1381 AGAGTTCTCTAATGTCGGAACCTCGGCCATGAAGGAAACGACAGCCTATATGATGACCGA 1440
 Db 1381 AGAGTTCTCTAATGTCGGAACCTCGGCCATGAAGGAAACGACAGCCTATATGATGACCGA 1440
 QY 1441 CATGATGAAACAGCTCTTGACTTATGGAATGGAACCTGACGAAATGCCCTATCTTGCCTGGCTCCC 1500
 Db 1441 CATGATGAAACAGCTCTTGACTTATGGAATGGAACCTGACGAAATGCCCTATCTTGCCTGGCTCCC 1500
 QY 1501 TCAGGCTGTAAACAGGAACTCTTAACCTATACAGACGAGGAAATTTGAAACCCACATCAA 1560
 Db 1501 TCAGGCTGTAAACAGGAACTCTTAACCTATACAGACGAGGAAATTTGAAACCCACATCAA 1560
 QY 1561 GACCTCTCAATTTGTAGCACCTGATGAACCTATTTGCTGGCTATACGCGTAAATATTTCAAT 1620
 Db 1561 GACCTCTCAATTTGTAGCACCTGATGAACCTATTTGCTGGCTATACGCGTAAATATTTCAAT 1620
 QY 1621 GCGTGTATGGACAGGCTATTCTAACCGTCTGACACACCTTTAGGCAATGGCTTACGGT 1680
 Db 1621 GCGTGTATGGACAGGCTATTCTAACCGTCTGACACACCTTTAGGCAATGGCTTACGGT 1680
 QY 1681 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACCTGCTGTAAGGAGCAATCCAGAAGA 1740
 Db 1681 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACCTGCTGTAAGGAGCAATCCAGAAGA 1740
 QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTTAAATATGGTCTCG 1800
 Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTTAAATATGGTCTCG 1800
 QY 1801 TTCTAGCTGGAATCACTACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 1860
 Db 1801 TTCTAGCTGGAATCACTACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 1860
 QY 1861 ATCAGATAGTTCAACTTTCACAGTCTACAGTCAACCTCCCAAGCACAATAATAGTAGAC 1920
 Db 1861 ATCAGATAGTTCAACTTTCACAGTCTACAGTCAACCTCCCAAGCACAATAATAGTAGAC 1920
 QY 1921 TACCAATCTCAATAATATACGCAACAATCAAAATCAACCCCTGATCAACAAATCAGAA 1980
 Db 1921 TACCAATCTCAATAATATACGCAACAATCAAAATCAACCCCTGATCAACAAATCAGAA 1980
 QY 1981 TCTTCAACCAAGCACAACCA 1999
 Db 1981 TCTTCAACCAAGCACAACCA 1999

QY 1681 CGCTGCCAAGTTTACCGCTCTATGATGACCTACCTGTCTGAGGAAGCAATCCAGAAGA 1740
 DB 8083 CGCTGCCAAGTTTACCGCTCTATGATGACCTACCTGTCTGAGGAAGCAATCCAGAAGA 8024
 QY 1741 TTGGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTAATTTAAAAATGGTCTCG 1800
 DB 8023 TTGGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTAATTTAAAAATGGTCTCG 7964
 QY 1801 TTCTAGTGGAACTCACCTGTCTCCACAACACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 DB 7963 TTCTAGTGGAACTCACCTGTCTCCACAACACCCCATCAACTGAAAGTTCAAGCTCATC 7904
 QY 1861 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACCACTCCAAAGCACAATAAATAGTACGAC 1920
 DB 7903 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACCACTCCAAAGCACAATAAATAGTACGAC 7844
 QY 1921 TACCATCTTCAATAATACCAACAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1980
 DB 7843 TACCATCTTCAATAATACCAACAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 7784
 QY 1981 TCCTCAACCAAGCACACCA 1999
 DB 7783 TCCTCAACCAAGCACACCA 7765

RESULT 5
 US-08-961-527-145/c
 ; Sequence 145, Application US/08961527
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/961,527
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: P3340P1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 145:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1071 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; US-08-961-527-145

Query Match 100.0%; Score 1999; DB 13; Length 10711;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAATAAAAAATCAATCTATGCTGACTTGGTGTCTGAAACGCCGCTCAA 60.

DB 9767 TAAATCTACGACAATAAAAAATCAATCTATGCTGACTTGGGTTCTGAAACGCCGCTCAA 9708
 QY 61 TGCCCAAGCTAATGATATTCCCAACAGATTGGTTAAGGAATCGTTTCTATCGAAGACCA 120
 DB 9707 TGCCCAAGCTAATGATATTCCCAACAGATTGGTTAAGGAATCGTTTCTATCGAAGACCA 9648
 QY 121 TCGCTTTCTCGACCAAGGGGATGATACCATCCGATATCTCTGGAGCTTTCTTGGCGAA 180
 DB 9647 TCGCTTTCTCGACCAAGGGGATGATACCATCCGATATCTCTGGAGCTTTCTTGGCGAA 9588
 QY 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCACTCTACCCCAACAGTTGATTAAGTTGAC 240
 DB 9587 TCTGCAAGCAATTCCTCCCAAGGTGGATCACTCTACCCCAACAGTTGATTAAGTTGAC 9528
 QY 241 TTACTTTTCAACTTCGACTTCCGACCAAGCTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
 DB 9527 TTACTTTTCAACTTCGACTTCCGACCAAGCTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 9468
 QY 301 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 360
 DB 9467 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 9408
 QY 361 GGTCTACATGCTTAATGGGAATGATGGAATGCAAGCAGAGCTCAAACTACTATGGTAA 420
 DB 9407 GGTCTACATGCTTAATGGGAATGATGGAATGCAAGCAGAGCTCAAACTACTATGGTAA 9348
 QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGAATGCTCAGCAGCACC 480
 DB 9347 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGAATGCTCAGCAGCACC 9288
 QY 481 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGAGCCGCGAAACTTGGTCTT 540
 DB 9287 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGAGCCGCGAAACTTGGTCTT 9228
 QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGATATGAGAAAGCAGCTCAATAC 600
 DB 9227 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGATATGAGAAAGCAGCTCAATAC 9168
 QY 601 ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGGA 660
 DB 9167 ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGGA 9108
 QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGGAAGCAAGGCTATAAAGCTACTCAC 720
 DB 9107 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGGAAGCAAGGCTATAAAGCTACTCAC 9048
 QY 721 AACTGGGATGGATGCTTACACAAATGATGACCAAGAGCTCAAAACATCTGTGGGATAT 780
 DB 9047 AACTGGGATGGATGCTTACACAAATGATGACCAAGAGCTCAAAACATCTGTGGGATAT 8988
 QY 781 TTACAATACAGAGCAATAGTTCCTTATCCAGACGATGAATTCGAAGTCTGCTTCTACCAT 840
 DB 8987 TTACAATACAGAGCAATAGTTCCTTATCCAGACGATGAATTCGAAGTCTGCTTCTACCAT 8928
 QY 841 TGTGTGATGTTTCTAACGGTAAAGTCAATGCCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 900
 DB 8927 TGTGTGATGTTTCTAACGGTAAAGTCAATGCCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 8868
 QY 901 TGTTCCTTCCGGAATTAACCAAGCAGTAGAACAACCCGGAGCTGGGGATCAACTATGAA 960
 DB 8867 TGTTCCTTCCGGAATTAACCAAGCAGTAGAACAACCCGGAGCTGGGGATCAACTATGAA 8808
 QY 961 ACCGATACAGACTATGCTTCCTGCTTGGAGTACGGTGTCTAGGATTCAGCTACTACTAT 1020
 DB 8807 ACCGATACAGACTATGCTTCCTGCTTGGAGTACGGTGTCTAGGATTCAGCTACTACTAT 8748
 QY 1021 CGTTACAGTACGCTTACCTGAGCAAAATCTCTGTTTATAAAGTGGGATAG 1080
 DB 8747 CGTTACAGTACGCTTACCTGAGCAAAATCTCTGTTTATAAAGTGGGATAG 8688
 QY 1081 GGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACAAATCGCGAAACGCTCCAGC 1140

Db	8687	GGGCTACTTTGGCAACATCACCTTGCATATACGCCCTGCACAATATCGCGGAACAGTCCACG	8628
Qy	1141	CGTGGAAACTCTAAACAAGTCGGACTCAACCGCGCCCAAGACTTTTCTTAATATGCTCTAGG	1200
Db	8627	CGTGGAAACTCTAAACAAGTCGGACTCAACCGCGCCCAAGACTTTCTTAATATGCTCTAGG	8568
Qy	1201	AATCGACTACCAAGTATTCTACTACTCAAAATGCCATTTCAAGTAAACACACCGAATCAGA	1260
Db	8567	AATCGACTACCAAGTATTCTACTACTCAAAATGCCATTTCAAGTAAACACACCGAATCAGA	8508
Qy	1261	CAAAAAATATGGAGCAAGTAGTCAAAAGATGGCTGCTCTACGCTGCCCTTTCGAAATGG	1320
Db	8507	CAAAAAATATGGAGCAAGTAGTCAAAAGATGGCTGCTCTACGCTGCCCTTTCGAAATGG	8448
Qy	1321	TGGAACCTTACTATAAAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGGAGTGA AAA	1380
Db	8447	TGGAACCTTACTATAAACC AATGTATATCCATAAAGTCGCTTTAGTGATGGGAGTGA AAA	8388
Qy	1381	AGAGTTCTCTAATGTGCGGAACCTGTGCCATGAAGGAAACGACGCCATATATGATGCCGA	1440
Db	8387	AGAGTTCTCTAATGTGCGGAACCTGTGCCATGAAGGAAACGACGCCATATATGATGCCGA	8328
Qy	1441	CATGATGAAAACAGTCTTGACTTATGGAACTGGAGCAAAATGCCCTATCTTCTTGGCTCCC	1500
Db	8327	CATGATGAAAACAGTCTTGACTTATGGAACTGGAGCAAAATGCCCTATCTTCTTGGCTCCC	8268
Qy	1501	TCAGGCTGGTAAAAACAGGAACCTCTAACTATACACAGCAGGAAATTTGAAAACACACATCAA	1560
Db	8267	TCAGGCTGGTAAAAACAGGAACCTCTAACTATACACAGCAGGAAATTTGAAAACACACATCAA	8208
Qy	1561	GACCTCTCAATTTGTAGCACCTGATGAAC TATTTGCTGGCTATACGCGTAAATATTCAAT	1620
Db	8207	GACCTCTCAATTTGTAGCACCTGATGAAC TATTTGCTGGCTATACGCGTAAATATTCAAT	8148
Qy	1621	GGCTGTATGACAGGCTATTCTTAACCGTCTGACACCACTTGTAGGCAATGGCCTTAGCGT	1680
Db	8147	GGCTGTATGACAGGCTATTCTTAACCGTCTGACACCACTTGTAGGCAATGGCCTTAGCGT	8088
Qy	1681	CGCTGCCAAAAGTTTACCGTCTATGATGACCTTACCTGTCTGAAGGAAGCAATCCAGAAGA	1740
Db	8087	CGCTGCCAAAAGTTTACCGTCTATGATGACCTTACCTGTCTGAAGGAAGCAATCCAGAAGA	8028
Qy	1741	TTGGAATATACAGAGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGGTGGCTCG	1800
Db	8027	TTGGAATATACAGAGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGGTGGCTCG	7968
Qy	1801	TTCTAGCTGGAAC TCACTTGCTTCCACACAAACCCCCCACTCAACTGAAAGTTCAAGCTCATC	1860
Db	7967	TTCTAGCTGGAAC TCACTTGCTTCCACACAAACCCCCCACTCAACTGAAAGTTCAAGCTCATC	7908
Qy	1861	ATCAGATAGTTC AACTTTCACAGTCTAGCTCAACCACTCCAAGCACAATATATAGTAGGAC	1920
Db	7907	ATCAGATAGTTC AACTTTCACAGTCTAGCTCAACCACTCCAAGCACAATATATAGTAGGAC	7848
Qy	1921	TACCAATTCCTTAACAATATACGCAACATCAAAATACAACCCCTGATCAACAAAAATCAGAA	1980
Db	7847	TACCAATTCCTTAACAATATACGCAACATCAAAATACAACCCCTGATCAACAAAAATCAGAA	7788
Qy	1981	TCCTCAACACAGCACACCA	1999
Db	7787	TCCTCAACACAGCACACCA	7769

```

RESULT 6
US-09-583-110-1312
; Sequence 1312, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26

```

; PRIOR APPLICATION NUMBER: US 09/107,433									
; PRIOR FILING DATE: 1998-06-30									
; PRIOR APPLICATION NUMBER: US 60/085,131									
; PRIOR FILING DATE: 1998-05-12									
; PRIOR APPLICATION NUMBER: US 60/051,553									
; PRIOR FILING DATE: 1997-07-02									
; NUMBER OF SEQ ID NOS: 5322									
; SEQ ID NO 1312									
; LENGTH: 2160									
; TYPE: DNA									
; ORGANISM: Streptococcus pneumoniae									
US-09-583-110-1312									
Query Match 99.4%; Score 1987.8; DB 22; Length 2160;									
Best Local Similarity 99.6%; Pred. No. 0;									
Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps									
Qy	1	TAA	AATCTACGACAAATAAAAAATCAACTATTGCTGTGACTTGGTTCCTGAACGCCGCGTCAA	600					
Db	159	taaa	ctacgacaaataaaatcaactcattgctgacttgggtctctgaaacgcgcgtcaa	218					
Qy	61	TGCC	CAAGCTAATGATATTCACACAGATTTGGTTAAGGCNAATCGTTTCATCGAAGACCA	120					
Db	219	tgcc	aaagctaagtatattccacagatttgggtaaggcaatcgtttctatcgaagaca	278					
Qy	121	TCG	CTTCTCGACCACAGGGGGATGTATACCATCCGTATCTCGGGAGCTTTCTTTGGCAA	180					
Db	279	tcgt	tcttcgaccacaaggggattgatacca tccgtatcctggtggagcttcttcgcgcaa	338					
Qy	181	TC	CTCAAAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCCAAACAGTTGATTAAAGTTGAC	240					
Db	339	ctc	gaagcaattccctccaagtggatcaactctcaactcaacagttgatlaagttgac	398					
Qy	241	TTA	CTTTTCAACTTCGACTTCGGACACGAGCTATTTCGTCTAAGGCTCAGGAAGCTTGTT	300					
Db	399	tta	cttccaacttcgacttccaccagactatttctcgttaaggctcaggaagcttggtt	458					
Qy	301	AGC	GATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360					
Db	459	agc	gattcagttagaacaaaaaagcaaccacaagaaatcttgacctactatataataa	518					
Qy	361	GGT	CTACATGCTTAATGGGAACATATGGAATGCAGACGAGCTCAAAAACACTACTATGGTAA	420					
Db	519	ggt	ctacatgcttaatggaaactatggaa tgcagacagcagctcaaaactactatggtaa	578					
Qy	421	AG	ACCTCAATAATTTAGTTTACCTCAGTTTAGCCTTGCTGGCTGGGAATGCCTCAGGCACC	480					
Db	579	aga	ctcaataaatttaagtttacctcagttgagcttgcgttggaatgcctcaggcacc	638					
Qy	481	AA	CCAAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCCGCCGAACCTTGCTCT	540					
Db	639	aaa	ccaata t gacccctattcacaatccagaagcagccccagaagccgcgaaacttggctct	698					
Qy	541	AT	CTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTAGTAGAAAAGCAGTCAATAC	600					
Db	699	atc	tgaatgaaaaatacaaggtcacatctctgtgaacagtagaagaagagtcacaaac	758					
Qy	601	AC	CAATTTACTGATGGACTCAAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA	660					
Db	759	acca	atlaactgctgaggtacaaagttcacaatcagcaagtaattacccttgcgttacatgga	818					
Qy	661	TA	ATTACTCAAGGAAGTCATCAATCAAGTTGAAGAAGAAACAGGCTATAACCTACTAC	720					
Db	819	taa	ttaactcaaggaagtcataatcaagt tgaagaagaacaaggcta taaactactcac	878					
Qy	721	AAC	TGGGATGGTCTCFACACAAATGTAGACCAAGAGCTCAAAAACATCTCTGGGATAT	780					
Db	879	aac	tgggatggga t gtcacacaatgcagaccaagaagctccaaaaaca tcttgggatatt	938					
Qy	781	TT	ACAATACAGAGCAATACGTTGCCTATCCAGACGATGAATTTGCAAGTCGCTTACCAT	840					
Db	939	tt	acaatfacagacaatatactttccctatccagacgcatgaatttcgaagtcgcttaccat	998					

QY 841 TGTTGATGTTTCTAACGGTAAGTCAATGCCCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 900
Db 999 tgttgatgtttctcaacggtaaagtcatgtcccaagctagagcagcgccatcagtcaagtaa 1058
QY 901 TGTTTCTCTCGGAATTAACCAAGCAGTACAGCAACAAACCGGAGTGGGATCAACTATGAA 960
Db 1059 tgtttctctcggaattaaaccaagcagtagaacaacacccgactggggtcaactatgaa 1118
QY 961 ACCGATCACAGACTATGCTCCTGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
Db 1119 accgatacacagactatgctcctgcttgggtacggtatctacgattcaactgctactat 1178
QY 1021 CGTTCAGATGAGCCCTATAACTACCCCTGGGCAAAATACTCCTGTTTATAACTGGGATAG 1080
Db 1179 cgttcacgatgagccctataactaccctgggacaaataccctgtttataactgggatag 1238
QY 1081 GGGCTACTTGGCAACATCACCTTGAATACGCCCTGCAACATCGGAAACGTCGCCAGC 1140
Db 1239 gggctacttggcaacatcaccttgcattgcacacccctgcacaaatcgcaaacgctccagc 1298
QY 1141 CGTGAACACTCTAAACAAGTCGAGCTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1200
Db 1299 cgtgaaactctaaacaaagtcgactcaacccgcgcaagacttctcctaagtgtctcgg 1358
QY 1201 AATCGACTACCCAAAGTATTCTACTTACTTCAATGCCATTTTCAAGTAAACACCAACGAATCAGA 1260
Db 1359 aatcgactacccaagtattcaactactcaaatgcatttcaagttaacacacacccgaatcaga 1418
QY 1261 CAAAAATATGGAGCAAGTAGTGAAGATGGCTGCTGTAGCTTACGCTTGCCTTTGCCAATGG 1320
Db 1419 caaaaaatagagcaagtagtgaagaatggctgtgtctacgctgccttggcaaatgg 1478
QY 1321 TGGAATCTACTATAAACCAACATGTATCCATAAAGTCGTCTTAGTGATGGGAGTGAATA 1380
Db 1479 tggaaactactataaaccaatgtatccataaagtcgtcttagtgatgggagtgaaaa 1538
QY 1381 AGAGTCTCTAATGTCGGAACGTCGTGCCATGAAGGAACACGACGCCTATATGATGACCGGA 1440
Db 1539 agagttctctaattgcggaactcgtgccatgaaggaaacgacgcctatatgatgaccca 1598
QY 1441 CATGATGAACACAGCTCTGACATATGGAATGGACCAATGCTATCTTCTGCTGGCTCCC 1500
Db 1599 catgagaacacagctctgacttatgtgaactggaactggaacatgctatcctgctggctccc 1658
QY 1501 TCAGGCTGTATAACAGGAACCTCTAACTATACAGACGAGGAAATGAAACACACATCAA 1560
Db 1659 tcaggctgtataaacaggaaaccttaactatacagacgaggaaaattgaaacacacatacaa 1718
QY 1561 GACCTCTCAATTTGTAGCACCTGATGAATATTTGCTGCTATACGCGTAAATATTTCAAT 1620
Db 1719 gacctctcaatttgtagcacctgatgaactatttggctgctatacgcgtataatattcaat 1778
QY 1621 GGCTGTATGACAGGCTATTCTAACGCTCTGACACACATTTAGGCAATGGCCTTACGGT 1680
Db 1779 ggcgtatggacaggctaattcaacgctctgacacacattgagcaatggccttaccgt 1838
QY 1681 CGCTGCAAGTTTACCGCTCTATGATGACCTACCTGCTGAGGAGCAATCCAGAA 1740
Db 1839 cgcgtccaaagtttaccgctctatgatgacctatctgtctggaaggaaagaatccagagga 1898
QY 1741 TTGGAATATACGAGGGGCTCTACAGAAATGGAGAAATTCGATTTTAAATAATGGTCTCG 1800
Db 1899 ttggaataaccagagggtctctacagaaatggagaattcgtatttaaaaaatggtgctcg 1958
QY 1801 TTCTAGTGGAACTACCTGCTCCACAAACCCCATCACTGAAGTTCAGCTCATT 1860
Db 1959 ttctacgttgaactaacctcgcctccacaacaccccccatcaactgaaagtccaagctcattc 2018
QY 1861 ATCAGATAGTTCAACTCTACAGTCTAGCTCAACACTCTCCAAAGCACAAATAATAGTACGAC 1920
Db 2019 atcagatagttcaacttcaagcttagctcaaccactccaaagcacaaaaataatagtagcagc 2078

QY 1921 TACCAATCTCTAACAAATAATACGCAACAATAACAACCCCTGATCAACAATAATCAGAA 1980
Db 2079 taccaatcttaacaataaataacgcaacaatacaaacctgatcaacaataatcagaa 2138
QY 1981 TCCTCAACCAAGCAGCAACCA 1999
Db 2139 tctcaaccagcacaacca 2157
RESULT 7
US-09-107-433-1102
; Sequence 1102, Application US/09107433
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIA
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2166 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1..2166
; SEQUENCE DESCRIPTION: SEQ ID NO: 1102:
US-09-107-433-1102
Query Match 99.4%; Score 1987.8; DB 15; Length 2166;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TAAATCTACGACAATAAAAAATCAACTCATTTGCTGACTTGGGTCTCTGAACGCGCGTCAA 60
Db 165 TAAATCTACGACAATAAAAAATCAACTCATTTGCTGACTTGGGTCTCTGAACGCGCGTCAA 224
QY 61 TGCCCCAAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
|||||

Db 225 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 284
QY 121 TCCTCTTCGACACAGGGGATGATACCATCCGATATCCCTGGGAGCTTTCTTCGCGAA 180
Db 285 TCCTCTTCGACACAGGGGATGATACCATCCGATATCCCTGGGAGCTTTCTTCGCGAA 344
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCCACACAGTTGATTAAAGTTGAC 240
Db 345 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCTCAACAGCTTGATTAAAGTTGAC 404
QY 241 TTACTTTTCACTTCGACCTTCGACACAGACTATTTCTCGTAAGGCTCAGAGAGCTTGGTT 300
Db 405 TTACTTTTCACTTCGACCTTCGACACAGACTATTTCTCGTAAGGCTCAGAGAGCTTGGTT 464
QY 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db 465 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 524
QY 361 GGTCTACATGTCTAATGGGAACCTATGGAATGCAGACAGAGCTCAAAACCTACTATGGTAA 420
Db 525 GGTCTACATGTCTAATGGGAACCTATGGAATGCAGACAGAGCTCAAAACCTACTATGGTAA 584
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGTAATGCCTCAGGACAC 480
Db 585 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGTAATGCCTCAGGACAC 644
QY 481 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCGGAAACTTGGCTTT 540
Db 645 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCGGAAACTTGGCTTT 704
QY 541 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 705 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 764
QY 601 ACCAATCTAGTGGACTACAAGTCTCAATCAGCAAGTAATACCCCTGCTTACATGGA 660
Db 765 ACCAATCTAGTGGGCTACAAGTCTCAATCAGCAAGTAATACCCCTGCTTACATGGA 824
QY 661 TAATTACCTCAGGAAGTATCAATCACTGAGGAGAAACAGGCTATAACTTACTAC 720
Db 825 TAATTACCTCAGGAAGTATCAATCACTGAGGAGAAACAGGCTATAACTTACTAC 884
QY 721 AACTGGGATGGATGCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
Db 885 AACTGGGATGGATGCTACACAAATAGACCAAGAGCTCAAAAACATCTGTGGGATAT 944
QY 781 TTACAATACAGAGATAGTTGCGCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
Db 945 TTACAATACAGAGATAGTTGCGCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 1004
QY 841 TGTGTGATGTTTCTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
Db 1005 TGTGTGATGTTTCTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 1064
QY 901 TGTTCCTCTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGGATCAACTATGAA 960
Db 1065 TGTTCCTCTCGGAATTAACCAAGCAGTAGAACAACCCGAGCTGGGGATCAACTATGAA 1124
QY 961 ACCGATCACAGACTATGCTCTGCGCTTGGAGTACGGTGTCTAGGATTCACCTGCTACTAT 1020
Db 1125 ACCGATCACAGACTATGCTCTGCGCTTGGAGTACGGTGTCTAGGATTCACCTGCTACTAT 1184
QY 1021 CGTTACAGTACGACCTATACCTACCTCGGACAAATACCTCTGTTTATAACTGGGATAG 1080
Db 1185 CGTTACAGTACGACCTATACCTACCTCGGACAAATACCTCTGTTTATAACTGGGATAG 1244
QY 1081 GGCTTACTTTGGCAACATCACCTTGCATACGCCCTGCAACAAATCGCGAAACGTCGCCAGC 1140
Db 1245 GGCTTACTTTGGCAACATCACCTTGCATACGCCCTGCAACAAATCGCGAAACGTCGCCAGC 1304
QY 1141 CGTGGAACTCTAACAAGGTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1200
Db 1305 CGTGGAACTCTAACAAGGTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1364

QY 1201 AATCGACTACCCAAAGTATTCACCTACTCAATGCCATTTTCAAGTAACACACCGAATCAGA 1360
Db 1365 AATCGACTACCCAAAGTATTCACCTACTCAATGCCATTTTCAAGTAACACACCGAATCAGA 1424
QY 1261 CAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGCTGTCTAGCGTGCCTTTTGCAAAATGG 1320
Db 1425 CAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGCTGTCTAGCGTGCCTTTTGCAAAATGG 1484
QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTAGTGTAGTGGAGTGAAGA 1380
Db 1485 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTAGTGTAGTGGAGTGAAGA 1544
QY 1381 AGAGTTCTCTAATGTCGGAACCTCGTCCCATGAAGGAAACGACAGCCTATATGATGACCGA 1440
Db 1545 AGAGTTCTCTAATGTCGGAACCTCGTCCCATGAAGGAAACGACAGCCTATATGATGACCGA 1604
QY 1441 CATGATGAAACAGTCTTGACCTTATGGAACCTGGACGAAATGCTTCTTGTGGCTCCC 1500
Db 1605 CATGATGAAACAGTCTTGACCTTATGGAACCTGGACGAAATGCTTCTTGTGGCTCCC 1664
QY 1501 TCAGGCTGTAAACAGGAACTCTAATCTATACAGACGAGGAAATTTGAAACACACATCAA 1560
Db 1665 TCAGGCTGTAAACAGGAACTCTAATCTATACAGACGAGGAAATTTGAAACACACATCAA 1724
QY 1561 GACCTCTCAATTTGTAGCACCTGATCAACTATTTGCTGGCTATACGCGTAAATATTTCAAT 1620
Db 1725 GACCTCTCAATTTGTAGCACCTGATCAACTATTTGCTGGCTATACGCGTAAATATTTCAAT 1784
QY 1621 GGCTGTATGACAGGCTATTTAACCGTCTGCACACACATTTGTAGGCAATGGCCTTACGGT 1680
Db 1785 GGCTGTATGACAGGCTATTTAACCGTCTGCACACACATTTGTAGGCAATGGCCTTACGGT 1844
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGTAAGGAGCAATCCAGAAGA 1740
Db 1845 CGCTGCCAAAGTTTACCGCTCTATGATGACCTATCTGCTGAAGGAGCAATCCAGAGA 1904
QY 1741 TTGGAATATACAGAGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1800
Db 1905 TTGGAATATACAGAGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1964
QY 1801 TTCTAGTGGAACTACCTGCTCTCCACAAACCCCATCAACTGAAAGTTTCAAGCTCATC 1860
Db 1965 TTCTAGTGGAACTACCTGCTCTCCACAAACCCCATCAACTGAAAGTTTCAAGCTCATC 2024
QY 1861 ATCAGATAGTTCAACTTCACAGTCTACAGCTCAACCCACTCCAGCACAATAATAGTAGCAG 1920
Db 2025 ATCAGATAGTTCAACTTCACAGTCTACAGCTCAACCCACTCCAGCACAATAATAGTAGCAG 2084
QY 1921 TACCAATCTTAACTAATACGCAACAATCAATAACAACCCCTGATCAACAAAATCAGAA 1980
Db 2085 TACCAATCTTAACTAATACGCAACAATCAATAACAACCCCTGATCAACAAAATCAGAA 2144
QY 1981 TCCTCAACGAGCACAACCA 1999
Db 2145 TCCTCAACGAGCACAACCA 2163

RESULT 8

PCT-US02-03987-9325
; Sequence 9325, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; PRIOR FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9325


```
Db 2139 tctcaaccagcacaacca 2157
|||||
RESULT 9
US-09-815-242-9325
; Sequence 9325, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9325
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
US-09-815-242-9325

Query Match 99.1%; Score 1981.4; DB 31; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAATAAAATCAACTCATTGCTGACTTGGGTTCTGAACCGCGGTCAA 60
Db 159 taaatctacgacaataaaatcaactcattgctgacttgggttctgaacgcgcgtcaa 218
QY 61 TGCCCAAGCTAATGATATTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 219 tgcccaagctaatgatattccacagatttggtaaggcaatcgtttctatcogaagacca 278
QY 121 TCGCTTTCGACACAGGGGATTGATACCATCCGTAFCCTGGGAGCTTCTTCGCGCAA 180
Db 279 tcgcttttcgacacaggggattgataccatcgtatcctcgtggagctttcttgcgcaa 338
QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCAACTCTCACCCCAACAGTTGATTAAGTTGAC 240
Db 339 tctgcaagcaattccctccaagggtggtacagctctcaactcaacagttgattaaattgac 398
QY 241 TTACTTTTCAACTCGACTTCGACACAGACTATTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 ttactttcaactcgcacttcgaccagactatttctcgttaaggctcaggaagcttggtt 458
QY 301 ACGGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 360
|||||
```



```
QY 1441 CATGATGAAACAGCTCTTGACTTATGGAACCTGGACGAAATGCCATCTTGGTGGCTCCC 1500
|||||
Db 1599 catgatgaaacagctcttgagttatggaactgacgaatgctatcttgcttgctccc 1658
QY 1501 TCAGGCTGTAAACAGGAACTCTAACTATACAGACGAGGAAATGAAACACATCAA 1560
|||||
Db 1659 tcaggctgttaaacaggaaacctctaactatacagacgaggaattgaaacacatcaa 1718
QY 1561 GACCTCTCAATTTGTAGCACCCTGATCAACTATTTGCTGGCTATACGCGTAAATATCAAT 1620
|||||
Db 1719 gacctctcaattttagcaccctgtagaactatttgcgtacacgcgtataatcaat 1778
QY 1621 GGCTGTATGACAGGCTATTTAAACGCTGTGACACCACTGTAGGCAATGGCCTTACGGT 1680
|||||
Db 1779 ggctgtatggacaggtattcttaaccgtctgcacacacttgtagcgaatggccttacgt 1838
QY 1681 CGCTGCAAAAGTTACCGCTCTATGATGACCTTACCTGTCTGAAGGAAGCAATCCAGAAG 1740
|||||
Db 1839 cgctgcgaagtttacgcctctatgatgacctacctgtctgaagggaagcaatccagagga 1898
QY 1741 TTGGAATATACCAAGAGGGGCTCTACAGAAATGGAGAAATTCGTATTTAAATAATGGTCTCG 1800
|||||
Db 1899 ttggaatataccagagggctctacagaaatggagaatcgtaatttaaaaaatggtgctcg 1958
QY 1801 TTCTAGCTGGAACTACCTGCTCCACAACAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
|||||
Db 1959 ttctacgtggagtcacctgctccacaacaaccccccatcaactgaaagtccaagtcctac 2018
QY 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAACGACACAATAATAGTAGCAG 1920
|||||
Db 2019 atcagatagtcaacttcacagctcagctcagctcaaccactccaagcacaataatagtagcag 2078
QY 1921 TACCAATCTCAACAATAATACGCAACAATAACAATCAACCCCTGATCAACAATAATCAGAA 1980
|||||
Db 2079 taccaatctcaacaataatcacgcaacaatacaatacaacccccctgatacaacaatacagaa 2138
QY 1981 TCCTCAACCAAGCACAACCA 1999
|||||
Db 2139 tctcaaccagcacacaacca 2157
```

```
RESULT 10
US-10-072-851-9325
; Sequence 9325, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits G
; TITLE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9325
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
US-10-072-851-9325

Query Match 99.1%; Score 1981.4; DB 37; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATTAATAAATCAACTCATCTGCTGACTTGGGTTCTGAACCGCGCTCAA 60
|||||
Db 159 taaatctcagcaataaaaaatcaactcatctgctgacttgggttctgaacgcgcgtcaa 218
QY 61 TGCCCAAGCTTAATGATATTCCCAACAGATTGGTTAAGGCAATCGTTTCTATCCAAGACCA 120
|||||
Db 219 tgcccaagcttaatgatatttcccaacagatttggttaagggaatcgtttctatcagaagacca 278
QY 121 TGCTTTCTCGACCAAGGGGATTCATACCATCCGTATCTCTGGGAGCTTTCTTGGCGAA 180
|||||
Db 279 tgcttcttcgacacaggggattgataccatccgtatcctctggagcttcttcttgcgcaa 338
QY 181 TCTGCAAAACCAATTCCTCCCAAGGTGGATCACTCTCACCCCAACAGCTGATTAAGTTGAC 240
|||||
Db 339 tctgcaaaacaaatctcctccaaggtggatcagctcctcactcaacagttgattgaattgac 398
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT 300
|||||
Db 399 ttactttcaactcgacttcgacacagactatttctgtaaggtcaggaagcttgggt 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
|||||
Db 459 agcgatttcagttagacaacaaagcaacaaagcaagaaatcttgacctactatataataa 518
QY 361 GGTCTACATCTCTAATGGGAACTATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA 420
|||||
Db 519 ggtctacatgtctaattgggaactatggaatgcagacagcagctcaaaactactatgtaa 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTTCCTGGCTGGAAATGCTCAGACACC 480
|||||
Db 579 agacctcaataatttaagtttacctcagtttagcttgcgttggatggatgctcaggcacc 638
QY 481 AAACCAATATGACCCCTTATTCACATCCAGAACGAGCCCAAGAGCCGCGAACTTGGTCTT 540
|||||
Db 639 aaaccaatagacccctattcacatccagaagcagcccaagacccgcgaacttggctt 698
QY 541 AUCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGATATGAGAAACGATCAATAC 600
|||||
Db 699 atctgaaatgaaaaatcaagggtcacatctctgctgacagatagagaagcagtcacat 758
QY 601 ACCAATTACTGTAGTGACTACAAAGTCTCAAAATCTCAATCAGCAAGTAATTAACCTTACATGGA 660
|||||
Db 759 accaattactgtatgggtcacaagtctcaaatcagcaagtaattacccctgcttaccatgga 818
QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGAAACAGGCTATAACCTACTCAC 720
|||||
Db 819 taattacctcaaggaaagtcacatcaatcaagttgaagaagaacaggtcataaactactcac 878
QY 721 AACTGGGATGGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGGGATAT 780
|||||
Db 879 aactgggatggatgctctacacaaatgtagaccaagagctcaaaaacatctgtgggatat 938
QY 781 TTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
|||||
Db 939 ttacaatacagacgaatacgttgcctatccagacagatgaattgcaagtgcgtctctaccat 998
QY 841 TGTGTGATGTTTCTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCGCATCAGTCAAGTAA 900
|||||
Db 999 tgttgtatgtttctaacgggtaaaagtcatttggccagctaggagcagcgcacatcagtaagtaa 1058
QY 901 TGTTTCTCTGGGAATTAACCAAGCACTAGAAACAAACCCGACTGGGGATCAACTATGAA 960
|||||
Db 1059 tgtttctctcggaattaaaccaagcagtagaacaacacccgcgactggggatcaactatgaa 1118
```

Qy	961	ACCGATCACAGACTATGCTCTCGCCTTGGAGTACGGGTGTACGATTCACTGCTACTAT	1020
Db	1119	accgatcacagactatgctcctgcttggagtacggtgtctacgagtccaatgcacat	1178
Qy	1021	CGTTCAGGATGAGCCCTATAACTTACCTCGGGACAATACTCCTGTTTAACTGGGATAG	1080
Db	1179	cgtttcaaga tgaagcccta taactacccctgggacaataaccctgtttataaac tgggatag	1238
Qy	1081	GGGCTACTTMTGGCAACATCACCTTGCATATACGCCCTGCAACAATCGGCAACGTCGCCAGC	1140
Db	1239	gggctactttggcaacatcaccttggaaatcacgacctgcaacaatcgcgaaacgtccccagc	1298
Qy	1141	CGTGGAAACTCTAAACAAGGTGGACTCAACCGCGCCGAAGACTTTCCTAATTGGTCTAGG	1200
Db	1299	cgtggaaactctaaacaaggtggactcaacgcgcgcgaagacttctctaagtgtctcgg	1358
Qy	1201	AATCGACTACCCAAAGTATTCACTTACTCAANTGCCATTTCAAAGTAACACAACCGAATCAGA	1260
Db	1359	aatcgactaccaagtattcaactactcaaatgccaattccaagtaacaacaaccgaatcaga	1418
Qy	1261	CAAAAAATATGGAGCAAGTAGTGAAAGATGCTGCTGCTTACGCTGCCTTTCGAAATGG	1320
Db	1419	caaaaaata tggagcaagtag tgaagaagatggctgtgcttacgctgaccttggcaaatg	1478
Qy	1321	TGGAACCTTACTATTAACCAATGTATATCCATTAAGTCCGCTTTTAGTGATGGGAGTGAAA	1380
Db	1479	tggaaacttacta caaaccaatgtataccaataagtcgtctttagtgatgggagtgaaa	1538
Qy	1381	AGAGTCTCTTAATGTGCGAACTCGTGCCATGAAGGAAACGACAGCCCTATATGATGACCGA	1440
Db	1539	agagtctctaa tgcgaaactcgtgccatgaagaagaacacacgcctatatgatgaccga	1598
Qy	1441	CATGATGAAAACAGCTTGTGACTTATGGAACTGGAGCAAAATGCCCTATCTTGTCTGGCTCCC	1500
Db	1599	catgatgaaaacagctctgagtta tggaaactggaagaatgcctatatctgtctggctccc	1658
Qy	1501	TCAGGCTGTGTAAACAGGAACTCTTAAGCTATACACGAGGAAATTTGAAACACACATCAA	1560
Db	1659	tcaggctgtgtaaaacaggaacctctaaactatacacagagagaaattgaaaaccacatcaa	1718
Qy	1561	GACCTCTCAATTTGTAGACCTGTAGTAAGTAATTTGCTGGCTATACGGGTAAATTTCAAT	1620
Db	1719	gacctctcaatttgtagcacttgaaactattgtgtgcttatcgcgtaaatattcaat	1778
Qy	1621	GGCTGTATGACAGGCTATTCTTAACCGTCTGACACCACTGTGAGCAATGSCCTACGGT	1680
Db	1779	ggctgtatggacaaggcta ttctaaacgtctgacacccacttgtaggaatggccctacggt	1838
Qy	1681	CGCTGCCAAAGTTTACCCTCTATGATGACCTTACCTGTCTGTGAAGGAAGCAATCCAGAAGA	1740
Db	1839	cgctgccaagt ttaccgtctctatgatgacctacctctgtgaaggagaacatccagagga	1898
Qy	1741	TTGGAATATACGAGAGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAATGGTGTCTCG	1800
Db	1899	ttggaaatataccagaagggtctacagaatggagaattcgtattttaaagtgggtcgt	1958
Qy	1801	TTCTACGTGGAATCACTGCTGCCAACAAACCCCACTCAACTGAAAGTTCAAGCTCATC	1860
Db	1959	ttctaaggagctcaactgctccaacaacaccccaatcaactgaagtccaagctcatc	2018
Qy	1861	ATCAGATAGTTCAACTTTCAGTCTAGTCAACCACTCCAAGGCACAATATATAGTAGCAC	1920
Db	2019	atcagatagtttcaacttcaagtctagctcaaccactccaagcacaataatagtagcac	2078
Qy	1921	TACCAATTCCTAACAAATTAATACGCAACAATTCAAATACAAACCCTCATCAACAAATACGAA	1980
Db	2079	taccaatlcctaaacaa taa tacygaacaatacctaaatacaaccctctgatcaacaaatcagaa	2138
Qy	1981	TCCTCAACGACGACAACCA 1999	
Db	2139	tcctcaaccagcacacaac 2157	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	99.1%;	Score 1981.4;	DB 45;	Length 10333;
	Best Local Similarity	99.4%;	Pred. No. 0;		
	Matches 1988;	Conservative	0; Mismatches	11; Indels	0; Gaps
Qy	1	TAAATCTACGACAATAAATAACATTCATTGCTGACTTGGTTCTGAAGCCCGGTCAA	60		
Db	9718	TAAATCTACGACAATAAATAACATTCATTGCTGACTTGGTTCTGAAGCCCGGTCAA	9659		
Qy	61	TGCCCAAGCTAATGATATTCACCAGATTTGGTTAAGCCAATCGTTCTATCGAAGACCA	120		
Db	9658	TGCCCAAGCTAATGATATTCACCAGATTTGGTTAAGCCAATCGTTCTATCGAAGACCA	9599		
Qy	121	TCGTTCTCTCCGACCACAGGGGGGATTGATACCATCCGTATCCTGGGAGGTTTCTTGCGCAA	180		
Db	9598	TCGTTCTCTCCGACCACAGGGGGGATTGATACCATCCGTATCCTGGGAGGTTTCTTGCGCAA	9539		
Qy	181	TCTGCAAGCAATTCCTCCGAAGTGGATCAACTCTCACCCAACAGTTGATTAAAGTTGAC	240		
Db	9538	TCTGCAAGCAATTCCTCCGAAGTGGATCAACTCTCACCCAACAGTTGATTAAAGTTGAC	9479		
Qy	241	TTACTTTTTCAACTTCGACTTCCGACCAGACATTTCTCGTAAGGCTCAGGAAGCTTGTTT	300		
Db	9478	TTACTTTTTCAACTTCGACTTCCGACCAGACATTTCTCGTAAGGCTCAGGAAGCTTGTTT	9419		

QY 301 AGCGATTGCTTAGAACAAAAGAACCAAGCAAGAAATCTTGGACCTACTATATATAATAA 360
Db 9418 AGCGATTGCTTAGAACAAAAGAACCAAGCAAGAAATCTTGGACCTACTATATAATAA 9359
QY 361 GGTCTACATGCTAATGGGAATATGGAAATGACAGACAGCTCAAAACTACTATGGTAA 420
Db 9358 GGTCTACATGCTAATGGGAATATGGAAATGACAGACAGCTCAAAACTACTATGGTAA 9299
QY 421 AGACCTCAATATATTAAGTTTACCTCAGTTAGCTTGGCTGGTGGAAATGCCCTCAGGCACC 480
Db 9298 AGACCTCAATATATTAAGTTTACCTCAGTTAGCTTGGCTGGTGGAAATGCCCTCAGGCACC 9239
QY 481 AAACCAATATGACCCCTATTTCATCTCAGAGACAGCCCAAGACCCCGGAAACTTGGTCTT 540
Db 9238 AAACCAATATGACCCCTATTTCATCTCAGAGACAGCCCAAGACCCCGGAAACTTGGTCTT 9179
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTCAACAGATGAGAAAGCAGTCAATAC 600
Db 9178 ATCTGAAATGAAATCAAGGCTACATCTCTGCTCAACAGATGAGAAAGCAGTCAATAC 9119
QY 601 ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 660
Db 9118 ACCAATTACTGATGGCTACAAAGTCTCAAAATCAGCAAGTAAATACCTGCTTACATGGA 9059
QY 661 TAAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAAGACAGCTATAACCTACTCAC 720
Db 9058 TAAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAAGACAGCTATAACCTACTCAC 8999
QY 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGACTCAAAACATCTGTGGGATAT 780
Db 8998 AACTGGGATGGATGCTACACAAATGTAGACCAAGACTCAAAACATCTGTGGGATAT 8939
QY 781 TTACAATACAGACGAATACGTTGGCTATCCAGACCATGAATTTGCAAGTCGGTCTACCAT 840
Db 8938 TTACAATACAGACGAATACGTTGGCTATCCAGACCATGAATTTGCAAGTCGGTCTACCAT 8879
QY 841 TGTGATGTTCTTAACGGTAAAGTCATTTGCCAGTGTAGAGACGCCATCAGTCAAGTAA 900
Db 8878 TGTGATGTTCTTAACGGTAAAGTCATTTGCCAGTGTAGAGACGCCATCAGTCAAGTAA 8819
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACCAACCGCAGCTGGGATCAACTATGAA 960
Db 8818 TGTTCCTTCGGAATTAACCAAGCAGTAGAACCAACCGCAGCTGGGATCAACTATGAA 8759
QY 961 ACCGATCACAGACTATGCTCTGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
Db 8758 ACCGATCACAGACTATGCTCTGCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 8699
QY 1021 GGTTCACATGAGCCCTATACTACCTCGGACAAATACCTCTGTTTATACTGGGATAG 1080
Db 8698 GGTTCACATGAGCCCTATACTACCTCGGACAAATACCTCTGTTTATACTGGGATAG 8639
QY 1081 GGGCTACTTTGGCAACATCACTTGCATACGCTGCAACAACTCGCAAGTCCAGC 1140
Db 8638 GGGCTACTTTGGCAACATCACTTGCATACGCTGCAACAACTCGCAAGTCCAGC 8579
QY 1141 CGTGAAACTCTAAACAGGTCGGAATCAACCGCGCCAGACTTTTCCCTAAATGGTCTAGG 1200
Db 8578 CGTGAAACTCTAAACAGGTCGGAATCAACCGCGCCAGACTTTTCCCTAAATGGTCTAGG 8519
QY 1201 AATGCTACCAAGTATTCTACTACTCAATAGCCATTTCAAGTAAACACCGGAATCAGA 1260
Db 8518 AATGCTACCAAGTATTCTACTACTCAATAGCCATTTCAAGTAAACACCGGAATCAGA 8459
QY 1261 CAAAAATATGAGCAACTAGTGAAGACTGCTGCTGCTACGCTGCTTGGCAATGG 1320
Db 8458 CAAAAATATGAGCAACTAGTGAAGACTGCTGCTGCTACGCTGCTTGGCAATGG 8399
QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGTTTGTAGTGGAGTGAAAA 1380
Db 8398 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGTTTGTAGTGGAGTGAAAA 8339
QY 1381 AGAGTTCTCTAATGTGGAACTCGTGCCATGAAGGAAACGACAGCCTATATGATGACCGA 1440

Db 8338 AGAGTTCTCTAATGTGGAACTCGTGCCATGAAGAAACGACAGCCTATATGATGACCGA 8279
QY 1441 CATGATGAAACAGCTTTGACTTATGGAACTGGAGCAATGCTTATCTTGTGCTGCC 1500
Db 8278 CATGATGAAACAGCTTTGACTTATGGAACTGGAGCAATGCTTATCTTGTGCTGCC 8219
QY 1501 TCAGCTGGTAAACAGGAACCTCTAACTATACAGACGAGGAAATTTGAAACCAATCAA 1560
Db 8218 TCAGCTGGTAAACAGGAACCTCTAACTATACAGACGAGGAAATTTGAAACCAATCAA 8159
QY 1561 GACCTCTCAATTTGTAGCACTGATGAACCTTTTCTGCTGGCTATACGCGTAAATATCAAT 1620
Db 8158 GACCTCTCAATTTGTAGCACTGATGAACCTTTTCTGCTGGCTATACGCGTAAATATCAAT 8099
QY 1621 GGCTGTATGGACAGGCTTATCTAAACGCTGTGACACCACTTTGAGGCAATGGCCTTACGGT 1680
Db 8098 GGCTGTATGGACAGGCTTATCTAAACGCTGTGACACCACTTTGAGGCAATGGCCTTACGGT 8039
QY 1681 CGCTGCCAAAGTTTACCCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA 1740
Db 8038 CGCTGCCAAAGTTTACCCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA 7979
QY 1741 TTGGAATATACAGAGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1800
Db 7978 TTGGAATATACAGAGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 7919
QY 1801 TTCTACGTGGAACTCACCTGCTCCACAAACCCCACTCAACTGAAAGTTCAAGCTCATC 1860
Db 7918 TTCTACGTGGAGCTCACCTGCTCCACAAACCCCACTCAACTGAAAGTTCAAGCTCATC 7859
QY 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCAGCAACAAATAATAGTACGAC 1920
Db 7858 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCAGCAACAAATAATAGTACGAC 7799
QY 1921 TACCAATCCTAACATAATACGCAACAAATCAAAATCACACCCCTGATCAACAAATCAGAA 1980
Db 7798 TACCAATCCTAACATAATACGCAACAAATCAAAATCACACCCCTGATCAACAAATCAGAA 7739
QY 1981 TCCTCAACACGACCAACCA 1999
Db 7738 TCCTCAACACGACCAACCA 7720

RESULT 12
US-60-068-175-596/c

; Sequence 596, Application US/60068175

; GENERAL INFORMATION:

; APPLICANT: Lagace, Robert E.

; APPLICANT: Corley, Neil C.

; APPLICANT: Russo, Frank D.

; APPLICANT: Hann, Amy L.

; APPLICANT: Heath, Joe D.

; APPLICANT: Finney, Gregory L.

; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE

; NUMBER OF SEQUENCES: 1175

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/60/068,175

; FILING DATE: HEREWITH

; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
 NAME: CERRONE, MICHAEL C.
 REGISTRATION NUMBER: 39,132
 REFERENCE/DOCKET NUMBER: PM-0009-2 P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (650) 855-0555
 TELEFAX: (650) 845-416
 INFORMATION FOR SEQ ID NO: 596:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9845 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: genomic DNA
 IMMEDIATE SOURCE:
 CLONE: SPN2C598
 US-60-068-175-596

Query Match 81.6%; Score 1631.2; DB 45; Length 9845;
 Best Local Similarity 88.8%; Pred. No. 0;
 Matches 1776; Conservative 0; Mismatches 223; Indels 1; Gaps 1;

QY 1 TAAATCTACGACATATAAATCAACTCATCTGCTGACCTTGGTTCTGAACGCCGCTCAA 60
 DB 9734 TAAATCTACGACATATAAATCAACTCATCTGCTGACCTTGGTTCTGAACGCCGCTCAA 9675

QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 DB 9674 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 9615

QY 121 TCGCTTCTTGACACAGGGGATTGATACCATCGTATCTCTGGAGCTTTCTTCGCGAA 180
 DB 9614 TCGCTTCTTGACACAGGGGATTGATTCATCGTATCTCTGGAGCTTTCTTCGCGAA 9555

QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCAACTCTCAACCAAGTTGATTAAGTTGAC 240
 DB 9554 TCTGCAAGCAATTCCTCCAGGTGGATCAACTCTCAACCAAGTTGATTAAGTTGAC 9495

QY 241 TTACTTTTCAACTTCGACCTCGACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT 300
 DB 9494 TTACTTTTCAACTTCGACCTCGACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT 9435

QY 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
 DB 9434 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 9375

QY 361 GGTCTACATGCTTAATGGAACTATGGAATGCAGACAGCTCAAAACTACTATGGTAA 420
 DB 9374 GGTCTACATGCTTAATGGAACTATGGAATGCAGACAGCTCAAAACTACTATGGTAA 9315

QY 421 AGACCTCAATATTAAGTTTACCTCAGTTAGCCTTGGCTGGAATGCCCTCAGGCACC 480
 DB 9314 AGACCTCAATATTAAGTTTACCTCAGTTAGCCTTGGCTGGAATGCCCTCAGGCACC 9255

QY 481 AAACCAATATGACCCCTATTACATCCAGAACAGCCCAAGACCCGCGAAACTTGGTCTT 540
 DB 9254 AAACCAATATGACCCCTATTACATCCAGAACAGCCCAAGACCCGCGAAACTTGGTCTT 9195

QY 541 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 600
 DB 9194 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 9135

QY 601 ACCAATTAAGTACGACATCAAGTCTCAATCAGCAAGTATTAACCTT-GCTTACATGG 659
 DB 9134 ACCAATTAAGTACGACATCAAGTCTCAATCAGCAAGTATTAACCTTGGCTTACATGG 9075

QY 660 ATAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAGAAACAGGCTATTAACCTTACCA 719
 DB 9074 ATAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAGAAACAGGCTATTAACCTTACCA 9015

QY 720 CAATCGGATGATGCTTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATA 779
 DB 7794 CAATCGGATGATGCTTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATA 7875

DB 9014 CTACTGGGATGGATGTTTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATA 8955

QY 780 TTTACAATACAGACGAATACGTTGCTTATCCAGACGATGAATTGCAAGTTCGCTTCTACCA 839

DB 8954 TCTACAATCCGATCAATACGCTCTTACCCCTGACGATGTTTGAAGTTCGATCTACGG 8895

QY 840 TTGTTGATGTTTCTAACGGTAAAGTCAATTGCCACGCTAGGAGCAGCCCATCAGTCAAGTA 899

DB 8894 TCGTAGATGTTTCAAAATGTAAGTCAATCGCCCAACTTGGAGCTCGTCAACCAAGCAATGA 8835

QY 900 ATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACACCGGACCTGGGATCACTATGA 959

DB 8834 ACGTTTCAATTTGGTACCAACCAAGCTGGAACAACCAATCGTACGTTGGTTCGTATGA 8775

QY 960 AACGATCACAGACTATGCTCTGCTGTTGGAGTACGGTGTCTACGATTCAACTGCTACTA 1019

DB 8774 AACCAATCACGATTTATGCACCTGCCATAGATAGCGGTGTTATGATTCCACTGCACTA 8715

QY 1020 TCGTTACGATGAGCCCTTATACTACCTGGGACAAATACCTCTCTTATTAACCTGGGATA 1079

DB 8714 TGGTTAATGATATCTTTATAACTATCCGGGAACAAGCACACCTCTCTACAACCTGGGATA 8655

QY 1080 GGGGTACTTTGGCAACATCACCCTTGCATAGCCCTCGAACAATCGGAAACGTCGCCAG 1139

DB 8654 GAGCATATTTCCGTAATATTAATCTCTGCAATATGCTCTTCAACAATCAGCAAAATGTCACAG 8595

QY 1140 CCGTGAACCTCTAAACAAGGTGCGACTCAACCGCGCAAGACTTTCTTAATGCTCTAG 1199

DB 8594 CCGTGAACCTCTTAAGGTGCGCTTAGATAGAGCTTAACACCTCTCTTAATGCTCTAG 8535

QY 1200 GAATCGACTACCCAAAGTATTCACCTACTCAAAATGCCATTTCAAGTAACACAACCGAATCAG 1259

DB 8534 GTATCGACTATCCAAAGCATGCAATTTGCAACGCCATTTCAAGTAACACAAGAACTCA 8475

QY 1260 ACAAAAATATGGAGCAAGTAGTGAAGAGATGGCTGCTGCTTACGCTGCTTTGCAATG 1319

DB 8474 ATAAACAATACGGAGCAAGTAGTGAAGAAATGGCTGCTGCTTATGCTTATGCTTTGCAATG 8415

QY 1320 GTGGAACTTACTATAAACAATGATATATCCATAAAGTCTGCTTTAGTATGGAGTGAAA 1379

DB 8414 GTGGAACTTACTATAAACAATGATATATCCATAAAGTCTGCTTTAGTATGGAGTGAAA 8355

QY 1380 AAGAGTCTCTAATGTCGGAACCTCGTGCATGGAAGAAACGACGCCCTATATGATGACCG 1439

DB 8354 AAGAGTCTCTAATGTCGGAACCTCGTGCATGGAAGAAACGACGCCCTATATGATGACCG 8295

QY 1440 ACATCATGAACACAGCTCTTGACTTATGGAACCTGGAGAAATGCCCTATCTTGGCTGGCTCC 1499

DB 8294 ACATCATGAACACAGCTCTTGACTTATGGAACCTGGAGAAATGCCCTATCTTGGCTGGCTCC 8235

QY 1500 CTCAGGCTGGTAAACAGGAACCTCTAACTATACAGACGAGGAAATTTGAAACACCATCA 1559

DB 8234 CTCAGGCTGGTAAACAGGAACCTCTAACTATACAGACGAGGAAATTTGAAACACCATCA 8175

QY 1560 AGACCTCTCAATTTGTAGCAGCTGATGAACCTATTTGCTGGCTATACGGCTAAATATTCAA 1619

DB 8174 AGAACACTGGCTATGCTCCAGATGAAATGTTTGGTGGTATATCTCTGTAAGTATCTTA 8115

QY 1620 TGGCTGATGACAGGCTTATCTAACCGCTGACACACCTTGTAGGCAATGSCCTTACGG 1679

DB 8114 TGGCTGATGACAGGCTTATCTGAATCGTTTAACTCTTATCGTTGGAGATGGTTTCCCTAG 8055

QY 1680 TCGCTGCCAAAGTTTACCGCTCTATGATGACCTACTGCTGAAGAAAGCAAGCAATCAGAG 1739

DB 8054 TTGACGCTTAAGTTTATCGCTCAATGATACGATATCTATCAGAAATACTCATCCAGAG 7995

QY 1740 ATTGGAATATACAGAGGGCTCTACAGAAATGGAAATTCGATTTAAAAATGTTGCTC 1799

DB 7994 ACTGGACGATGCCAGACGAGCTTTTCAGAAACGGGAAATTTGATTTCAAAAATGAGCTC 7935

QY 1800 GTTCTACGTGGAACCTGCTGCTCCACAAACACCCCTCACTCAAGTTCAAGCTCAT 1859

DB 7934 GCGCAATATGGACTGAACCCCTCTACTCAACAATCTCTCAACAGCTGAAAGTTCAAGCTCAT 7875

QY 1860 CATGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCAAAGCACAATAATAGTACGA 1919
Db 7874 CATGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCAAAGCACAATAATAGTACGA 7815
QY 1920 CTACCAATCCTAACCAATATAGCGCAACCAATCAATATACACCCCTGATCAACAAATACGA 1979
Db 7814 CTACCAATCCTAACCAATATAGCGCAACCAATCAATATACACCCCTGATCAACAAATACGA 7755
QY 1980 ATCCCTCAACGACGACCAACA 1999
Db 7754 ATCCCTCAACGACGACCAACA 7735

RESULT 13
US-08-116-541-3
; Sequence 3, Application US/08116541
; GENERAL INFORMATION:
; APPLICANT: Masure, H. Robert
; APPLICANT: Pearce, Barbara J.
; APPLICANT: Tounanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/116.541
; FILING DATE: 19930901
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 960 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6
; IMMEDIATE SOURCE:
; CLONE: SPRU42
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..960

Query Match 47.28; Score 944.2; DB 5; Length 960;
Best Local Similarity 99.78; Pred. No. 1e-256;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TAAATCTACGACAATAAAATCAACTATGCTGACTTGGGTTCTGAACGCCGGTCAA 60

Db 12 TAAATCTACGACAATAAAATCAACTATGCTGACTTGGGTTCTGAACGCCGGTCAA 71
QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 120
Db 72 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGACCA 131
QY 121 TCGCTTCTTCGACCACAGGGGATTGATACCATCGGTATCTCTGGGAGCTTCTTCTGGCAA 180
Db 132 TCGCTTCTTCGACCACAGGGGATTGATACCATCGGTATCTCTGGGAGCTTCTTCTGGCAA 191
QY 181 TCTGCAAGCAATTCCTCCAAAGTGGATTAACCTCTCAACCAAGTGTGATTAAGTTGAC 240
Db 192 TCTGCAAGCAATTCCTCCAAAGTGGATTAACCTCTCAACCAAGTGTGATTAAGTTGAC 251
QY 241 TTACTTTTCAACTTCGACTTCGACCAAGTGGATTAACCTCTCAACCAAGTGTGATTAAGTTGAC 300
Db 252 TTACTTTTCAACTTCGACTTCGACCAAGTGGATTAACCTCTCAACCAAGTGTGATTAAGTTGAC 311
QY 301 AGCGATTTCAGTTAGAACCAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 360
Db 312 AGCGATTTCAGTTAGAACCAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 371
QY 361 GGTCTACATGCTTAATGGGAATATGGAAATGCAGACAGCTCAAAACTACTATATGGTAA 420
Db 372 GGTCTACATGCTTAATGGGAATATGGAAATGCAGACAGCTCAAAACTACTATATGGTAA 431
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCCTTGGCTGGAAATGCCTCAGGCACC 480
Db 432 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCCTTGGCTGGAAATGCCTCAGGCACC 491
QY 481 AAACCAATATGACCCCTATTACATCCAGAACGACGCCAGCAAGCTTGGTCTT 540
Db 492 AAACCAATATGACCCCTATTACATCCAGAACGACGCCAGCAAGCTTGGTCTT 551
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 552 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 611
QY 601 ACCAATTACTGATGACTACAAAGTCTCAAAATCAGCAAGTAAATACCCTGCTTACATGA 660
Db 612 ACCAATTACTGATGATGCTACAAAGTCTCAAAATCAGCAAGTAAATACCCTGCTTACATGA 671
QY 661 TAAATACCTCAAGGAAGTCAATCAAGTTGAAGAAGAAACAGGCTATAACCTACTCAC 720
Db 672 TAAATACCTCAAGGAAGTCAATCAAGTTGAAGAAGAAACAGGCTATAACCTACTCAC 731
QY 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGTCAAAACATCTGTGGGATAT 780
Db 732 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGTCAAAACATCTGTGGGATAT 791
QY 781 TTACAATACAGACGAATACGTTGCCATATCCAGACCATGAATTCGAAGTCGCTTCTACCAT 840
Db 792 TTACAATACAGACGAATACGTTGCCATATCCAGACCATGAATTCGAAGTCGCTTCTACCAT 851
QY 841 TGTGATGTTTCTAACGGTAAAGTCAATTCGCCAGCTAGGACGACGCCATCAGTCAAGTAA 900
Db 852 TGTGATGTTTCTAACGGTAAAGTCAATTCGCCAGCTAGGACGACGCCATCAGTCAAGTAA 911
QY 901 TGTGATGTTTCTAACGGTAAAGTCAATTCGCCAGCTAGGACGACGCCATCAGTCAAGTAA 949
Db 912 TGTGATGTTTCTAACGGTAAAGTCAATTCGCCAGCTAGGACGACGCCATCAGTCAAGTAA 960

RESULT 14
PCT-US97-14436-139
; Sequence 139, Application PC/TUS9714436
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard

APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
SEQUENCE CHARACTERISTICS: POLYPEPTIDES AND THEIR USES
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
PCT-US97-14436-139

Query Match 40.5%; Score 808.6; DB 1; Length 2172;

Best Local Similarity 99.4%; Pred. No. 5.1e-218;

Matches 822; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1174 CCCCAAGACTTTCCCTAAATGGTCTAGGAATCGACTACCCAACTATTCTACTCTCAAAATGC 1233
DB 1 CCCCAAGACTTCCCCTAAATGGTCTAGGAATCGACTACCCAACTATTCTACTCTCAAAATGC 60
QY 1234 CATTTCAGTAACACACCGCAATCAGACAAAAATATGAGCAAGTAGTGAAGATGCG 1293
DB 61 CATTTCAGTAACACACCGCAATCAGACAAAAATATGAGCAAGTAGTGAAGATGCG 120
QY 1294 TGCTGCTTACGTCGCTTTGCAATGGTGGAACTTACTATTAACCAATGTATATCCATAA 1353
DB 121 TGCTGCTTACGTCGCTTTGCAATGGTGGAACTTACTATTAACCAATGTATATCCATAA 180
QY 1354 AGTCGCTTTAGTGTAGGAGTGAAGAGTTCTTAATGTCGGAACCTCGTGCATGAA 1413
DB 181 AGTCGCTTTAGTGTAGGAGTGAAGAGTTCTTAATGTCGGAACCTCGTGCATGAA 240
QY 1414 GGAACGACGACGCTATATGATGACCGACATGATGAAAAACAGTCTTGAACCTTGAACCTGG 1473
DB 241 GGAACGACGACGCTATATGATGACCGACATGATGAAAAACAGTCTTGAACCTTGAACCTGG 300
QY 1474 ACG-NAATGCCCTATCTTGGTGGCTCCCTCAGGCTGGTAAACAGGAACTCTTAACATA 1532
DB 301 ACGAAATGCCCTATCTTGGTGGCTCCCTCAGGCTGGTAAACAGGAACTCTTAACATA 360
QY 1533 CAGACGAGGAATTTGAAACACCATCAAGACTCTCAATTTGTAGCACTGTATGAACAT 1592
DB 361 CAGACGAGGAATTTGAAACACCATCAAGACTCTCAATTTGTAGCACTGTATGAACAT 420

QY 1593 TTGCTGGCTATACGGCTAAATTAATCAATGGCTGTATGAGACAGGCTATTCTTAACCGTCTGA 1652
DB 421 TTGCTGGCTATACGGCTAAATTAATCAATGGCTGTATGAGACAGGCTATTCTTAACCGTCTGA 480
QY 1653 CACCACTTGTAGGCAATGCGCTTACGGTCTGCGTCCAAAGTTTACCGCTCTATGATGACCT 1712
DB 481 CACCACTTGTAGGCAATGCGCTTACGGTCTGCGTCCAAAGTTTACCGCTCTATGATGACCT 540
QY 1713 ACCTGCTCTGAAGGAACCAATCCAGAGGATTTGGAATATACCGAGGGGCTCTACAGAAATG 1772
DB 541 ACCTGCTCTGAAGGAACCAATCCAGAGGATTTGGAATATACCGAGGGGCTCTACAGAAATG 600
QY 1773 GAGAAATCGTATTAAAAATGGTCTGCTTACGTGGAACCTCACTGCTCCACAAAC 1832
DB 601 GAGAAATCGTATTAAAAATGGTCTGCTTACGTGGAACCTCACTGCTCCACAAAC 660
QY 1833 CCCCATCAACTGAAAGTTCAAGCTCATCATGATAGTTCAACTTCACAGTCTAGCTCAA 1892
DB 661 CCCCATCAACTGAAAGTTCAAGCTCATCATGATAGTTCAACTTCACAGTCTAGCTCAA 720
QY 1893 CCACCTCCAAAGCACAAATAATAGTAGTACGACTACCAATCTTAACAATAATACGCAACAATCAA 1952
DB 721 CCACCTCCAAAGCACAAATAATAGTAGTACGACTACCAATCTTAACAATAATACGCAACAATCAA 780
QY 1953 ATACAACCCCTGATCAACAATAATCAAGATCCTCAACACGACCAACCA 1999
DB 781 ATACAACCCCTGATCAACAATAATCAAGATCCTCAACACGACCAACCA 827
RESULT 15
US-08-911-503-139
; Sequence 139, Application US/08911503
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
; POLYPEPTIDES AND THEIR USES
; NUMBER OF SEQUENCES: 713
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911.503
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,022
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50533
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-911-503-139

Query Match	40.5%	Score 808.6;	DB 13;	Length 2172;
Best Local Similarity	99.4%	Pred. No. 5.1e-218;		
Matches 822;	Conservative 0;	Mismatches 4;	Indels 1;	Gaps 1;
QY 1174 CGCCAGAGCTTCCCTAAATGGTCTAGGAATCGACTACCCAAAGTATTTCAGTACTCAATGC 1233				
Db 1 CGCCAGAGCTTCCCTAAATGGTCTAGGAATCGACTACCCAAAGTATTTCAGTACTCAATGC 60				
QY 1234 CATTTCAGTAACACAAACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAAAGATGGC 1293				
Db 61 CATTTCAGTAACACAAACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAAAGATGGC 120				
QY 1294 TGCCTGTACGCTGCCCTTTGGCAATGGTGAACCTTACTATATAACCAATGTATATCCATAA 1353				
Db 121 TGCCTGTACGCTGCCCTTTGGCAATGGTGAACCTTACTATATAACCAATGTATATCCATAA 180				
QY 1354 AGTCGCTTTTAGTGTGGAGTGAAAAGAGTCTCTAATGTCGGAACTCGTGCCATGAA 1413				
Db 181 AGTCGCTTTTAGTGTGGAGTGAAAAGAGTCTCTAATGTCGGAACTCGTGCCATGAA 240				
QY 1414 GGAACGACAGCCTATATGATGACCCGACATGATGAAAACAGTCTTGACTTATGGAACCTGG 1473				
Db 241 GGAACGACAGCCTATATGATGACCCGACATGATGAAAACAGTCTTGACTTATGGAACCTGG 300				
QY 1474 ACG-AAATGCCCTATCTTGGTCCCTCAGGCTGGTAAAAACAGGAACCTCTAAGTATA 1532				
Db 301 ACGAAAATGCCCTATCTTGGTCCCTCAGGCTGGTAAAAACAGGAACCTCTAAGTATA 360				
QY 1533 CAGACGAGGAATGGAACACCATCAACACCTCTCAATTTGTAGCACCTGATGAACCTAT 1592				
Db 361 CAGACGAGGAATGGAACACCATCAACACCTCTCAATTTGTAGCACCTGATGAACCTAT 420				
QY 1593 TTGCTGGCTATACGGGTAAATATTCAATGGCTGTATGGACAGGCTATTCTAACCGTCTGA 1652				
Db 421 TTGCTGGCTATACGGGTAAATATTCAATGGCTGTATGGACAGGCTATTCTAACCGTCTGA 480				
QY 1653 CACCACTTGTAGGCAATGGCCTTACGGTCGCTGCCAAAGTTTACCGCTCTATGATGACCT 1712				
Db 481 CACCACTTGTAGGCAATGGCCTTACGGTCGCTGCCAAAGTTTACCGCTCTATGATGACCT 540				
QY 1713 ACCTGCTGAAGGAACCAATCCAGAGATTGGAATATCCAGAGGGGCTCTACAGAAATG 1772				
Db 541 ACCTGCTGAAGGAACCAATCCAGAGATTGGAATATCCAGAGGGGCTCTACAGAAATG 600				
QY 1773 GAGAAATTCGATTAAAAATGGTCTGCTTCTACGTGGAACCTCACCTGCTCCACAAACAAC 1832				
Db 601 GAGAAATTCGATTAAAAATGGTCTGCTTCTACGTGGAACCTCACCTGCTCCACAAACAAC 660				
QY 1833 CCCCATCAACTGAAAGTTTCAAGCTCATCATCAGATAGTTTCAACTTCAAGTCTAGCTCAA 1892				
Db 661 CCCCATCAACTGAAAGTTTCAAGCTCATCATCAGATAGTTTCAACTTCAAGTCTAGCTCAA 720				
QY 1893 CCACCTCCAGCAACAATAATAGTACGACTACCAATCCCTAACAATATATAGCAACATCAA 1952				
Db 721 CCACCTCCAGCAACAATAATAGTACGACTACCAATCCCTAACAATATATAGCAACATCAA 780				
QY 1953 ATACACCCCTGATCAACAAAAATCAGAAATCCTCAACGACGACCAACCA 1999				
Db 781 ATACACCCCTGATCAACAAAAATCAGAAATCCTCAACGACGACCAACCA 827				

Search completed: June 13, 2002, 14:43:46
Job time: 7603 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 12:42:33 ; Search time 279.06 Seconds
(without alignments)
10068.144 Million cell updates/sec

Title: US-08-961-083-1

Perfect score: 1999
Sequence: 1 TAAATCTACGACATAAAAA.....ATCTCAACGACACCA 1999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 997330 seqs, 702755440 residues

Total number of hits satisfying chosen parameters: 1994660

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	1999	100.0	1999	4	US-08-961-083-1		Sequence 1, Appl
2	117.2	5.9	9100	6	US-10-121-120-27		Sequence 27, Appl
3	59.2	3.0	2370	5	US-09-540-2098-3719		Sequence 3719, Ap
4	55.2	2.8	428	5	US-09-673-476-206		Sequence 206, App
5	51.8	2.6	363	5	US-09-673-476-626		Sequence 626, App
6	40.6	2.0	223	5	US-09-673-476-438		Sequence 438, App
7	40	2.0	610	6	US-10-027-632-5967		Sequence 5967, Ap
8	37.4	1.9	430	6	US-10-027-632-40693		Sequence 40693, A
9	37.4	1.9	242	5	US-09-789-189-1228		Sequence 1228, Ap
10	37.2	1.9	1611	7	US-60-360-039-46474		Sequence 46474, A
11	36.6	1.8	679	6	US-10-027-632-213202		Sequence 213202,
12	36.6	1.8	2409	7	US-60-360-039-46149		Sequence 46149, A
13	36.4	1.8	473	5	US-09-673-476-671		Sequence 671, App
14	36	1.8	689	6	US-10-027-632-148124		Sequence 148124,
15	35.8	1.8	348	5	US-09-673-476-99		Sequence 99, Appl
16	35.6	1.8	2373	5	US-09-540-2098-4136		Sequence 4136, Ap
17	35.4	1.8	8045	5	US-09-913-878A-1		Sequence 1, Appl
18	35	1.8	510	6	US-10-027-632-82144		Sequence 82144, A
19	35	1.8	510	6	US-10-027-632-309095		Sequence 309095,
20	35	1.8	132205	5	US-09-578-519B-25		Sequence 25, Appl
21	34.8	1.7	3435	1	PCR-US02-10624-12		Sequence 12, Appl
22	34.6	1.7	1247	6	US-10-027-632-254816		Sequence 254816,
23	34.6	1.7	1247	6	US-10-027-632-254817		Sequence 254817,
24	34.6	1.7	1247	6	US-10-027-632-254818		Sequence 254818,
25	34.6	1.7	1247	6	US-10-027-632-254819		Sequence 254819,
26	34.6	1.7	1247	6	US-10-027-632-254820		Sequence 254820,

27	34.6	1.7	2466	5	US-09-769-744A-121	Sequence 121, App	
28	34.6	1.7	2486	6	US-10-104-047-1358	Sequence 1358, App	
c	29	34.2	1.7	622	6	US-10-027-632-235624	Sequence 235624, A
30	34.2	1.7	747	7	US-60-360-039-32748	Sequence 32748, A	
31	34.2	1.7	2463	7	US-60-360-039-26589	Sequence 26589, A	
c	32	34.2	1.7	14803	6	US-10-105-299-11424	Sequence 11424, A
33	34	1.7	585	6	US-10-027-632-219980	Sequence 219980, App	
34	34	1.7	653	6	US-10-123-155-438	Sequence 438, App	
35	34	1.7	653	6	US-10-137-871-438	Sequence 438, App	
36	34	1.7	653	6	US-10-141-761-438	Sequence 438, App	
37	34	1.7	653	6	US-10-140-864-438	Sequence 438, App	
38	34	1.7	653	6	US-10-140-923-438	Sequence 438, App	
39	34	1.7	653	6	US-10-141-756-438	Sequence 438, App	
40	34	1.7	653	6	US-10-141-759-438	Sequence 438, App	
41	34	1.7	653	6	US-10-140-472-438	Sequence 438, App	
42	34	1.7	653	6	US-10-140-805-438	Sequence 438, App	
43	34	1.7	653	6	US-10-142-885-438	Sequence 438, App	
44	34	1.7	653	6	US-10-146-731-438	Sequence 438, App	
45	34	1.7	653	6	US-10-142-426-438	Sequence 438, App	

ALIGNMENTS

RESULT 1
US-08-961-083-1
: Sequence 1, Application US/08961083
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE: 30-Oct-1997
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: <Unknown>
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1999 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-961-083-1

Query Match 100.0%; Score 1999; DB 4; Length 1999;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAATCTACGACATAAAAAATCACTCAATGCTGGTCTGTAACGCCGCTCAA 60
|||||

Db 1 TAAATCTACGACAAATAAATAAATCAATCTATGCTGACTTGGGTCTTGAACGCCGCGTCAA 60
 QY 61 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 Db 61 TGCCCAAGCTAATGATATATCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
 QY 121 TCGCTTTCTCGACCAAGGGGATGATACCATCCGTATCCCTGGGAGCTTTCTTGGCGAA 180
 Db 121 TCGCTTTCTCGACCAAGGGGATGATACCATCCGTATCCCTGGGAGCTTTCTTGGCGAA 180
 QY 181 TCTGCAAAACAATTCCTCCCAAGGTGGATCAACTCTCACCCCAACAGTTGATTAAGTTGAC 240
 Db 181 TCTGCAAAACAATTCCTCCCAAGGTGGATCAACTCTCACCCCAACAGTTGATTAAGTTGAC 240
 QY 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT 300
 Db 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT 300
 QY 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
 Db 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
 QY 361 GGTCTACATGCTAATGGGAATATGGAATGCAAGCAGAGCTCAAAACTACTATGGTAA 420
 Db 361 GGTCTACATGCTAATGGGAATATGGAATGCAAGCAGAGCTCAAAACTACTATGGTAA 420
 QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGGCTGGTGAATGCCCTCAGGCACC 480
 Db 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGGCTGGTGAATGCCCTCAGGCACC 480
 QY 481 AAACCAATATGACCCCTATTCACATCCAGAAGCAGCCCAAGCAGCCGCAAGCTTGGTCTT 540
 Db 481 AAACCAATATGACCCCTATTCACATCCAGAAGCAGCCCAAGCAGCCGCAAGCTTGGTCTT 540
 QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCAGTCAATAC 600
 Db 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCAGTCAATAC 600
 QY 601 ACCAATTAAGTGGTACTGAAAGTCTCAAAATCAGCAAGTAAATACCCCTGCTTACATGA 660
 Db 601 ACCAATTAAGTGGTACTGAAAGTCTCAAAATCAGCAAGTAAATACCCCTGCTTACATGA 660
 QY 661 TAATTAACCTCAAGGAAGTCAATCAAGTTGAAGAAAGCAGGCTATAAAGCTACTAC 720
 Db 661 TAATTAACCTCAAGGAAGTCAATCAAGTTGAAGAAAGCAGGCTATAAAGCTACTAC 720
 QY 721 AACTGGGATGGATGCTACACAAATGTCAGCAAGTAAATGCAAGTAAATGCAAGTAAAT 780
 Db 721 AACTGGGATGGATGCTACACAAATGTCAGCAAGTAAATGCAAGTAAATGCAAGTAAAT 780
 QY 781 TTACAATACAGCAATACGTTGCCTATCCAGACGATGAATGCAAGTAAATGCAAGTAAAT 840
 Db 781 TTACAATACAGCAATACGTTGCCTATCCAGACGATGAATGCAAGTAAATGCAAGTAAAT 840
 QY 841 TGTGATGTTTCAAGGTAAGTCAATTCGCCAGCTAGGACGACGCGCATCAGTCAAGTAA 900
 Db 841 TGTGATGTTTCAAGGTAAGTCAATTCGCCAGCTAGGACGACGCGCATCAGTCAAGTAA 900
 QY 901 TGTGATGTTTCAAGGTAAGTCAATTCGCCAGCTAGGACGACGCGCATCAGTCAAGTAA 960
 Db 901 TGTGATGTTTCAAGGTAAGTCAATTCGCCAGCTAGGACGACGCGCATCAGTCAAGTAA 960
 QY 961 ACCGATACAGACTATGCTCTCCCTGGGATGAGTACGCTGCTACGATTCAGCTACTAT 1020
 Db 961 ACCGATACAGACTATGCTCTCCCTGGGATGAGTACGCTGCTACGATTCAGCTACTAT 1020
 QY 1021 CGTTCCAGTATGACCTATTAACCTACCTGGGCAATAATCTCTGTTTATAAATGGGATAG 1080
 Db 1021 CGTTCCAGTATGACCTATTAACCTACCTGGGCAATAATCTCTGTTTATAAATGGGATAG 1080
 QY 1081 GGGCTACTTTGGCAATCAGCTTGAATACGCTGCAACCAATCGGAAAGCTCCGACG 1140
 Db 1081 GGGCTACTTTGGCAATCAGCTTGAATACGCTGCAACCAATCGGAAAGCTCCGACG 1140

QY 1141 CGTGGAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCTCTAAATGGTCTAGG 1200
 Db 1141 CGTGGAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCTCTAAATGGTCTAGG 1200
 QY 1201 AATCGACTACCCCAAGTATTCACCTACTACTCAATGCCATTTCAAGTAACACACCAATCAGA 1260
 Db 1201 AATCGACTACCCCAAGTATTCACCTACTACTCAATGCCATTTCAAGTAACACACCAATCAGA 1260
 QY 1261 CAATAAATATGAGCAAGTATGCAAAAGATGGCTGCTGCTTACGCTGGCTTTTCCAATGG 1320
 Db 1261 CAATAAATATGAGCAAGTATGCAAAAGATGGCTGCTGCTTACGCTGGCTTTTCCAATGG 1320
 QY 1321 TGAACCTTACTATAAACCAATATGATATCAATAAAGTCGCTTTAGTAGTGGGAGTGAAGA 1380
 Db 1321 TGAACCTTACTATAAACCAATATGATATCAATAAAGTCGCTTTAGTAGTGGGAGTGAAGA 1380
 QY 1381 AGAGTTCTCTAATGTCGGAACCTCTAATGTAAGTAAGGAAGACGACGCTATATGATGACCGA 1440
 Db 1381 AGAGTTCTCTAATGTCGGAACCTCTAATGTAAGTAAGGAAGACGACGCTATATGATGACCGA 1440
 QY 1441 CATGATGAAACAGTCTTGCATTTATGGAATGGAACCTGACGCTATGCTTGGCTCCC 1500
 Db 1441 CATGATGAAACAGTCTTGCATTTATGGAATGGAACCTGACGCTATGCTTGGCTCCC 1500
 QY 1501 TCAGGCTGTAAACAGGAACTCTAATGTAAGTAAGGAAGACGACGCTATGATGATGACCGA 1560
 Db 1501 TCAGGCTGTAAACAGGAACTCTAATGTAAGTAAGGAAGACGACGCTATGATGATGACCGA 1560
 QY 1561 GACCTCTCAATTTGTAGCAGCTGATGAACCTTATGCTGCTATACGCTGATGATGATGATGAT 1620
 Db 1561 GACCTCTCAATTTGTAGCAGCTGATGAACCTTATGCTGCTATACGCTGATGATGATGATGAT 1620
 QY 1621 GCTGTATGAGCAGGCTATTTAAACGCTCTGACACCACTTGTAGGCAATGGCTTACGGT 1680
 Db 1621 GCTGTATGAGCAGGCTATTTAAACGCTCTGACACCACTTGTAGGCAATGGCTTACGGT 1680
 QY 1681 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACCTGCTGAGGAAGCAATCCAGGAAGA 1740
 Db 1681 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACCTGCTGAGGAAGCAATCCAGGAAGA 1740
 QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATATTTAAAAATGGTCTCG 1800
 Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATATTTAAAAATGGTCTCG 1800
 QY 1801 TTCTACGTGAACTACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 Db 1801 TTCTACGTGAACTACCTGCTCCACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
 QY 1861 ATCAGATGTTCAACTTACAGTCTACAGTCTACAGTCTACAGTCTACAGTCTACAGTCTACAG 1920
 Db 1861 ATCAGATGTTCAACTTACAGTCTACAGTCTACAGTCTACAGTCTACAGTCTACAGTCTACAG 1920
 QY 1921 TACCATCTCAATAATACGCAACCAATCAATAACCCCTGATCAACAAATCAGAA 1980
 Db 1921 TACCATCTCAATAATACGCAACCAATCAATAACCCCTGATCAACAAATCAGAA 1980
 QY 1981 TCCTCAACCAAGCACAACCA 1999
 Db 1981 TCCTCAACCAAGCACAACCA 1999

RESULT 2

US-10-121-120-27/c

; Sequence 27, Application US/10121120

; GENERAL INFORMATION:

; APPLICANT: Bergeron, Michel G.

; APPLICANT: Ouellette, Marc

; APPLICANT: Roy, Paul H.

; TITLE OF INVENTION: Specific and Universal Probes and Amplification

; TITLE OF INVENTION: Primers

; TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial
 ; Pathogens and Antibiotic Resistance Genes from Clinical

```

; TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro
; FILE REFERENCE: 12287.31
; CURRENT APPLICATION NUMBER: US/10/121,120
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/452,599
; PRIOR FILING DATE: 1999-12-01
; PRIOR APPLICATION NUMBER: 08/304,732
; PRIOR FILING DATE: 1994-09-12
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 9100
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-10-121-120-27

```

Query Match	5.9%	Score 117.2	DB 6	Length 9100
Best Local Similarity	50.3%	Pred. No. 1.5e-22		
Matches 349	Conservative 0	Mismatches 333	Indels 12	Gaps 2
QY 22	TCAACTCATTGCTGACATTGGTTCCTGACGCCGGTCAATGCCCAAGCTAATGATATCC 81			
Db 3498	TAAATTAATTGGCGAAGTGGGTAGCAACGCCGATTCCAGTGAATTAGCCGATGTC 3439			
QY 82	CACAGATTGGTTAAGCAATCGTTCTATCGAAGACCATCGCTCTCTCGACACACAGGG 141			
Db 3438	ACACGCTTAATTGACGCATTTTAGCGCGGAAGACAGTCGTTTTCAGTATCATCAGG 3379			
QY 142	GATTGATACCATCCGTATCTCGGGAGCTTTCTTCGGCAATCTGCAAA---GCAATTCCCT 198			
Db 3378	ATTAGACCCATTGGCATTTGCCGTGCATTTGTCGCACTGAGTAATGCGGTGCATC 3319			
QY 199	CCRAGGTGGATCAACTCTCACCCAAACAGTTGATTAAAGTGNCTTACHTTTCACATTCCGAC 258			
Db 33:8	ACAAGGCGAAGTACGATTACTCAACAATTAGCCGTAACTTTTTCTTA-----AC 3268			
QY 259	TTCCGACGACACTATTCTCGTAAGGCTCAGGAAGCTTGGTTAGCGATTCAAGTTAGAACA 318			
Db 3267	CTCAGAAAAACCATTTATCGTAAGCTCGTGAAGCCGTGCTTCGGGTAGNAATCGAAAA 3208			
QY 319	AAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAAGTCTACATGTCTAATGG 378			
Db 3207	TACTCTCAACAAACAGAAATATTAGAGCTTTATTAAACAAAATCTTTTAGCTATCG 3148			
QY 379	GAACTATGGATGCAGACAGCAGCTCAAACTACTATGTTAAAGACCTCAATATTTTAAG 438			
Db 3147	TTCTATATGGTGTGTCAGCGCGACGACAAACCTATTTCGGTAAATCATTTGAATGAATTGAC 3088			
QY 439	TTTACCTCAGTTAGCCCTTGCTGGTAATGCCCTCAGGCACCAAAACCAATATGACCCCTA 498			
Db 3087	CTTATCGGAATGCGATTATTGCTGGTTTACCTTAAGCACCTTCAACAATGAACCCGCT 3028			
QY 499	TTCAATCCAGAGCAGCCCAAGACCCGGAACCTTGGTCTTATCTGAAATGAAAAATCA 558			
Db 3027	TTATTCTTTAAAAAGTTTCAGAAAGAACCCGCAATGTGGTGTGCTAAGCCGTATTAGATGA 2968			
QY 559	AGGCTACATCTCTGCTCAACAGATATGAAAGCAGTCAATACACCAATTACTGATGGACT 618			
Db 2967	AAATACATCAGCAAGAAATATGATGTCATTGAAAGCGCGATTGTGGCGAGCTA 2908			
QY 619	ACAAGTCTCAAAATGACCAAGTATTACCTCGCTTACATGGATTAATTACCTCAAGGAAGT 678			
Db 2907	TCAGGGCGCAAAATTTGAATTTCCAGCCGATTAATGTCACATGAATGGTGGCTCAAGAAAT 2848			
QY 679	CATCAATCAAGTTGAAGAGAAACAGCGCTATAAC 712			
Db 2847	GGTGCCTCGTTTGTGGCGAAGAAATGCTTACACC 2814			

RESULT 3
US-09-540-209B-3719
: Sequence 3719. Application US/09540209B

```

; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 3719
; LENGTH: 2370
; TYPE: DNA
; ORGANISM: B.fragilis
; US-09-540-209B-3719

```

[illegible]

RESULT 4
US-09-673-476-206
; Sequence 206, Application US/09673476
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; APPLICANT: BUCHRIESER-BROSCH, ROLAND
; APPLICANT: GORDON, STEPHEN
; APPLICANT: BILLAULT, ALAIN
; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
; TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
; TITLE OF INVENTION: MYCOBACTERIA.
; FILE REFERENCE: 05394. 0011-00000

```

Best Local Similarity 52.6%; Pred. No. 0.00013;
Matches 113; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 326 ACCAAGCAAGAAATCTTGACCTACTATATAAATAAGGTCTACATGTCTAATGGAACTAT 385
   ||| ||||| ||||| ||| ||| ||||| ||| ||| |||
Db 26 acaaaatctgaatcctgacccgatacttgaacctggtctcgctcggaataaactcgctc 85

QY 386 GGAATGCAGACAGCGCTCAAACTACTATGTTGTAAGACCTCAATAATTTAACTTTACCT 445
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 86 ggqgtcaggacgcggcgcaacgctacttcggcatcaacgcgccgacctgaattggcag 145

QY 446 CAGTTACCTTGCTGGTGGAAATGCCCTCAGGCACCAACCAATATGACCCCTATTACAT 505
   ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 146 caagcgcgctgctggcgcgcatggtgcatcgaccagcagctcaaccgctacccaac 205

QY 506 CCAGAAGCAGCCGAAGACCCGCCGAAACTTGTGTT 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 ccqgacgcgctggtggccggcggaacgtggtcct 240

RESULT 6
US-09-673-476-438
; Sequence 438, Application US/09673476
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; APPLICANT: BUCHRIESER-BROSCH, ROLAND
; APPLICANT: GORDON, STEPHEN
; APPLICANT: BILLAULT, ALAIN
; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
; TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF
; FILE REFERENCE: 05394.0011-00000
; CURRENT APPLICATION NUMBER: US/09/673,476
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: PCT/IB99/00740
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 09/060,756
; PRIOR FILING DATE: 1998-04-16
; NUMBER OF SEQ ID NOS: 743
; SOFTWARE: Patent In Ver. 2.2
; SEQ ID NO 438
; LENGTH: 223
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (136)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (158)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (203)
; OTHER INFORMATION: a, t, c or g
US-09-673-476-438

Query Match 2.0%; Score 40.6; DB 5; Length 223;
Best Local Similarity 48.6%; Pred. No. 0.16;
Matches 106; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 281 AAGCTCAGGAAGCTTGTTAGCGATTCACTTATAGACAAAGAACGACCAAGCAAGAATC 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 aagcttcgagatccggatggcactcagctggacaagaccttcacaaatcgtgaatc 65

QY 341 TTGACCTACTATATAATAGGTCTACATGTCTAATGGGAACATATGGAATGCAGACGCA 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 ctgaccgcgactgaacctggtctcgctcggaataaactcgctcgcgcgagacgcg 125

QY 401 GCTCAAAACTACTATGTAAGAACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGCTG 460

```

```
Db 126 gcgcaacgctnctcgcatcaacgcgtcgactgaattggcagcaaggcgctgctg 185
QY 461 GCTGGAATGCTCAGGCAACCAACCAATATGACCCCTA 498
Db 186 gcgcgcatggtgcaatcaccagcagctcaaccgta 223

RESULT 7
US-10-027-632-5967/c
; Sequence 5967, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5967
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-5967

Query Match 2.0%; Score 40; DB 6; Length 610;
Best Local Similarity 55.1%; Pred. No. 0.34; Indels 0; Gaps 0;
Matches 76; Conservative 1; Mismatches 61;

QY 1845 AAGTTCAAGCTCATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCCCAAGCA 1904
Db 410 ACAGYGAARCTCCATCTCAATACTACTACTACTACTACTACTACTACTACTACTACTA 351
QY 1905 CAAATAATAGTAGGACTCAATCTTAACAATAATACGCAACAATCAATACAAACCCCTG 1964
Db 350 CTACTACTACTACTAATAATAATAATAATAATAATGAGAGCTAAACATATTTACCTG 291
QY 1965 ATCAACAAATCAGATC 1982
Db 290 TAGGCACTATCTCAATC 273

RESULT 8
US-10-027-632-40693/c
; Sequence 40693, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40693
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(430)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-40693

Query Match 1.9%; Score 38; DB 6; Length 430;
Best Local Similarity 60.8%; Pred. No. 1.1; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 40;

QY 319 AAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAAATAGGTCTCATGTCTATGG 378
Db 392 AAAAGCAAGTGGGAAAAAACAACCAACAATAAAATGAGAGCTGAATTTTAAACGG 333
QY 379 GAACATATGAATCCAGACAGCAGCTCAAAACTACTATGGTAA 420
Db 332 AAGCAATGGAATTCAGAAGCAATAGAAATAATATTTTAAAA 291

RESULT 9
US-09-789-189-1228
; Sequence 1228, Application US/09789189
; GENERAL INFORMATION:
; APPLICANT: Lelias, Jean-Michel
; TITLE OF INVENTION: Human Polynucleotides and Polypeptides
; FILE REFERENCE: 25436/1720
; CURRENT APPLICATION NUMBER: US/09/789,189
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/183452
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 2005
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1228
; LENGTH: 242
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(242)
; OTHER INFORMATION: "n" at position 33 refers to any of G, A, T or C
US-09-789-189-1228

Query Match 1.9%; Score 37.4; DB 5; Length 242;
Best Local Similarity 54.4%; Pred. No. 1.3; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 62;

QY 308 CAGTTAGACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAAATAGGTCTAC 367
Db 32 cnatcaaaaaaaaaaaaaaaaaaaaaaacaaattatcatatacaaacataataaaaaa 91
QY 368 ATGCTTAATGGGAACATATGGAATGCAGACAGCAGCTCAAAACTACTATGTTAAAGACCTC 427
Db 92 caatccatctcaaacataaaaaataaataaattctaaactcctattataaacatctt 151
QY 428 AATAATTTAAGTTTAC 443
```

Db 152 tataataaactaac 167

RESULT 10
US-60-360-039-46474
; Sequence 46474, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46474
; LENGTH: 1611
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-60-360-039-46474

Query Match 1.9%; Score 37.2; DB 7; Length 1611;
Best Local Similarity 53.4%; Pred. No. 3;
Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 746 GTAGACCAAGAGCTCAAAACATCTGTGGATATTTACAATACAGACGAATACGTTGCC 805
Db 829 ggagatgaagtggtgtatccttgcctatttatattcaagaatacaaaaacgttga 888
QY 806 TATCCAGACGAATGTAAGTCAAGTCGTTCTACCATTTGATGTTTCTACGGTAAAGTC 865
Db 889 aagccaattgatgattttacgtcttttgcactactgtctctagtctatggttaaagtc 948
QY 866 ATGCCAGCTAGGACGCGCATCA 891
Db 949 atggccaaagtagaaggattccatca 974

RESULT 11
US-10-027-632-213202/c
; Sequence 213202, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 213202
; LENGTH: 679
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-213202

Query Match 1.8%; Score 36.6; DB 6; Length 679;
Best Local Similarity 49.2%; Pred. No. 3.2;
Matches 96; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 683 AATCAAGTTGAAGAGAAACAGGCTATAACCTACTCACAACCTGGGATGGATGCTTACACA 742
Db 568 AAACATAAGAAATAAGAAATCAGATTTTAAATATGTTAGGTTCAAACTCAAAAACA 509
QY 743 AATGTAGACCAAGAGCTCAAAACATCTGTGGGATATTTACAATACAGACGAATACGTT 802
Db 508 AAGAGAAATACAGATCAACACATCAATCACTATAAAATTTATTAAAAATTAAGAAGTT 449
QY 803 GCCTATCCAGAGGATGAATTCGAAGTCGCTTCTACCATTTGTCATGTTTCTACGCTAAA 862
Db 448 CTTTCTCCAGATCTTTTAAAGTTGTTTTGGTTTGAATACTACTTCTGACTTTCAA 389
QY 863 GTCAITGCCAGCTA 877
Db 388 ACAATTCAAAACATA 374

RESULT 12
US-60-360-039-46149
; Sequence 46149, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46149
; LENGTH: 2409
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-60-360-039-46149

Query Match 1.8%; Score 36.6; DB 7; Length 2409;
Best Local Similarity 52.3%; Pred. No. 5.2;
Matches 81; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 673 GGAAGTCATCAATCAAGTTGAAGAGAAACAGGCTATAACCTACTCACAACCTGGGATGA 732
Db 243 ggaagtaataaacacagaagaagaagaagaacatcaacaaaagcggaatga 302
QY 733 TGTCTACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATATTTACAATACAGA 792
Db 303 tgatgacgatgatgaatgaagaagaagaagaagaagaagaatgatgatga 362
QY 793 CGAATACGTTGCCTATCCAGACGATGAATTCGAAG 827
Db 363 tgatgacgacgatgatgaagaagaagaagaaga 397

RESULT 13
US-09-673-476-671/c
; Sequence 671, Application US/09673476
; GENERAL INFORMATION:
; APPLICANT: COLE, STEWART
; APPLICANT: BUCHRIESER-BROSCH, ROLAND
; APPLICANT: GORDON, STEPHEN
; APPLICANT: BILLAULT, ALAIN
; TITLE OF INVENTION: A METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST
; TITLE OF INVENTION: FROM THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED
; TITLE OF INVENTION: DNA LIBRARY. APPLICATION TO THE DETECTION OF

Db 201 TTGCGGACACAGCGGCGTGGACTGGAAGGGCACCCCTG 163

Search completed: June 13, 2002, 14:47:52
Job time: 7519 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 08:39:33 ; Search time 3635.99 Seconds
(without alignments)
11505.022 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTACGACAAATRAAA.....ATCCTCAACGACACAACCA 1999

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 15

Total number of hits satisfying chosen parameters: 37854

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_hgt_hum:*

31: em_hgt_inv:*

32: em_hgt_other:*

33: em_hgtg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999	100.0	10711	6	AR120238 Sequence
2	1999	100.0	11334	6	BD003825 Polynucle
3	1999	100.0	11334	1	AE007349 Streptoco
4	1795	89.8	2160	1	AF210745 Streptoco
5	1591	79.6	2157	1	SPPONAS
6	1489	74.5	2160	1	AF210746
7	1438	71.9	2157	1	SPPONAS
8	1438	71.9	2160	6	AX111315 Sequence
9	1438	71.9	3378	1	STRPONA
10	1438	71.9	10313	1	AE008414 Streptoco
11	1198	59.9	1249	6	AX111327 Sequence
12	889	44.5	1195	1	AB006874 Streptoco
13	835	41.8	1195	1	AB006868 Streptoco
14	796	39.8	960	6	AR084885 Sequence
15	774	38.7	1826	1	SPPONAS
16	705	35.3	2154	1	AF210748 Streptoco
17	671	33.6	930	1	AF046233 Streptoco
18	620	31.0	930	1	AF046235 Streptoco
19	620	31.0	930	1	AF046237 Streptoco
20	620	31.0	930	6	AX111320 Sequence
21	590	29.5	1199	6	AX110281 Sequence
22	567	28.4	1229	6	AX110284 Sequence
23	551	27.6	930	1	AF046236 Streptoco
24	461	23.1	930	1	AF046234 Streptoco
25	436	21.8	2160	1	AF210747 Streptoco
26	436	21.8	30457	2	SPNEU1914
27	432	21.6	2160	1	AF210752 Streptoco
28	396	19.8	2160	6	AX111314 Sequence
29	396	19.8	2172	1	SPCS1091A
30	396	19.8	2172	1	SPCS1111A
31	387	19.4	489	1	SPN403978
32	387	19.4	2157	1	SPPONAS
33	318	15.9	2157	1	SPPONAS
34	318	15.9	2157	1	SPPONAS
35	318	15.9	2157	1	SPPONAS
36	318	15.9	2157	1	SPPONAS
37	318	15.9	2157	6	AX111407 Sequence
38	318	15.9	2160	1	AF139883 Streptoco
39	318	15.9	2160	1	AF139884 Streptoco
40	318	15.9	2160	1	AF139885 Streptoco
41	318	15.9	2160	1	AF139886 Streptoco
42	318	15.9	2160	1	AF139887 Streptoco
43	318	15.9	2160	1	AF139888 Streptoco
44	318	15.9	2160	1	AF139889 Streptoco
45	318	15.9	2160	1	AF139890 Streptoco
46	318	15.9	2160	1	AF159448 Streptoco
47	318	15.9	2160	1	AF210749 Streptoco
48	318	15.9	2160	1	AF210750 Streptoco
49	318	15.9	2160	1	AF210751 Streptoco
50	318	15.9	2160	6	AX110436 Sequence
51	318	15.9	2160	6	AX111316 Sequence
52	298	14.9	1195	1	AB006869 Streptoco
53	298	14.9	1195	1	AB006872 Streptoco
54	298	14.9	1195	1	AB006873 Streptoco
55	245	12.3	1195	1	AB006879 Streptoco
56	228	11.4	1195	1	AB006877 Streptoco
57	228	11.4	1195	6	AX111319 Sequence
58	224	11.2	567	1	SPPBP1AA
59	209	10.5	489	1	SPN403977
60	205	10.3	1195	1	AB006870 Streptoco
61	205	10.3	1195	1	AB006878 Streptoco
62	203	10.2	930	1	AF046232 Streptoco
63	197	9.9	930	1	AF046231 Streptoco
64	178	8.9	1212	6	AX111323 Sequence
65	178	8.9	1222	6	AX110283 Sequence
66	175	8.8	1220	6	AX110280 Sequence
67	159	8.0	1216	6	AX111329 Sequence
68	159	8.0	2157	1	SPPONAS
69	153	7.7	1195	1	AB006876 Streptoco
70	148	7.4	1195	1	AB006871 Streptoco
71	131	6.6	489	1	SPN403979
72	129	6.5	1201	6	AX110279 Sequence
73	127	6.4	1212	6	AX110272 Sequence

74	127	6.4	1223	6	AX110275	AX110275 Sequence	147	20	1.0	113688	9	AC005301	AC005301 Homo sapi
75	126	6.3	1211	6	AX110282	AX110282 Sequence	c 148	20	1.0	118969	8	H0212B02	AL442007 Oryza sat
76	126	6.3	1218	6	AX110274	AX110274 Sequence	c 149	20	1.0	119770	8	AC008697	AC008697 Homo sapi
77	122	6.1	1213	6	AX110273	AX110273 Sequence	150	20	1.0	124823	9	AC007064	AC007064 Homo sapi
78	122	6.1	1214	6	AX110276	AX110276 Sequence	151	20	1.0	139634	2	AP004223	AP004223 Oryza sat
79	121	6.1	1207	6	AX110278	AX110278 Sequence	152	20	1.0	141371	2	AC084087	AC084087 Homo sapi
80	120	6.1	1212	6	AX110271	AX110271 Sequence	153	20	1.0	145684	2	AC016588	AC016588 Homo sapi
81	116	5.8	489	1	SPN403980	AJ403980 Streptococ	c 154	20	1.0	146120	9	AC024942	AC024942 Homo sapi
82	115	5.8	1242	6	AX111324	AX111324 Sequence	c 155	20	1.0	148183	2	AP004672	AP004672 Oryza sat
83	115	5.5	1223	6	AX1110277	AX1110277 Sequence	c 156	20	1.0	151024	2	AC016855	AC016855 Homo sapi
84	104	5.2	1225	6	AX111325	AX111325 Sequence	c 157	20	1.0	158344	10	AL133401	AL133401 Mouse DNA
85	97	4.9	930	1	AF046230	AF046230 Streptococ	158	20	1.0	158467	9	AC011314	AC011314 Homo sapi
86	97	4.9	930	6	AX111317	AX111317 Sequence	159	20	1.0	161735	9	AC068633	AC068633 Homo sapi
87	90	4.5	782	6	AX111331	AX111331 Sequence	160	20	1.0	161788	9	AC009314	AC009314 Homo sapi
88	90	4.5	813	6	AX110915	AX110915 Sequence	161	20	1.0	165725	2	OSJN00105	AL066692 Oryza sat
89	77	3.9	489	1	SPN403976	AJ403976 Streptococ	162	20	1.0	166204	2	AC026763	AC026763 Homo sapi
90	73	3.7	1225	6	AX110285	AX110285 Sequence	163	20	1.0	167063	2	AL353758	AL353758 Homo sapi
91	44	2.2	489	1	SPN403974	AJ403974 Streptococ	c 164	20	1.0	168720	2	AC078781	AC078781 Homo sapi
92	44	2.2	489	1	SPN403981	AJ403981 Streptococ	165	20	1.0	169340	9	AP001148	AP001148 Homo sapi
93	44	2.2	930	1	AF046238	AF046238 Streptococ	c 166	20	1.0	170232	9	AL135903	AL135903 Human DNA
94	44	2.2	930	1	AF387163	AF387163 Streptococ	c 167	20	1.0	170751	2	HSBA536C5	AL121987 Homo sapi
95	44	2.2	930	6	AX111318	AX111318 Sequence	168	20	1.0	175933	2	AC096595	AC096595 Rattus no
96	44	2.2	1582	1	SMAJ2290	AJ002290 Streptococ	169	20	1.0	176236	2	AC026344	AC026344 Homo sapi
97	34	1.7	2181	1	SMI295856	AJ295856 Streptococ	c 170	20	1.0	177564	9	AP002799	AP002799 Homo sapi
98	29	1.5	1201	1	AB006875	AB006875 Streptococ	171	20	1.0	180319	2	AC016918	AC016918 Homo sapi
99	26	1.3	930	1	AY072702	AY072702 Streptococ	c 172	20	1.0	187985	2	AL611932	AL611932 Homo sapi
100	25	1.3	25	6	AX110484	AX110484 Sequence	c 173	20	1.0	188078	9	AC008045	AC008045 Homo sapi
101	23	1.2	33	6	ARI20351	ARI20351 Sequence	c 174	20	1.0	194334	2	AC046140	AC046140 Homo sapi
102	23	1.2	2528	1	STRPONNA	M90528 Streptococ	c 175	20	1.0	194674	2	AC021168	AC021168 Homo sapi
103	22	1.1	22	6	AX110459	AX110459 Sequence	176	20	1.0	195349	2	AC006705	AC006705 Caenorhab
104	22	1.1	4534	1	AX111282	AX111282 Sequence	177	20	1.0	197110	9	AC104306	AC104306 Homo sapi
105	22	1.1	4534	1	EFA302065	AJ302065 Enterococ	c 178	20	1.0	197110	17	AC010742	AC010742 Homo sapi
106	22	1.1	189452	2	AL663110	Mus muscu	c 179	20	1.0	201065	2	AL627236	AL627236 Mus muscu
107	21	1.1	138317	9	AP003481	AP003481 Homo sapi	c 180	20	1.0	201215	9	AC019084	AC019084 Homo sapi
108	21	1.1	152816	2	AC079739	AC079739 Homo sapi	c 181	20	1.0	214187	9	AC008703	AC008703 Homo sapi
109	21	1.1	162608	2	AC009478	AC009478 Homo sapi	c 182	20	1.0	234270	2	AL627342	AL627342 Mus muscu
110	21	1.1	167322	2	AC016788	AC016788 Homo sapi	183	19	1.0	19	6	A83833	A83833 Sequence 11
111	21	1.1	184767	2	AC084363	AC084363 Homo sapi	184	19	1.0	19	6	AY037187	AY037187 Arabidops
112	21	1.1	302387	2	AC026348	AC026348 Homo sapi	c 185	19	1.0	382	8	G49014	G49014 SHGC-78177
113	20	1.0	20	6	AX110392	AX110392 Sequence	186	19	1.0	465	11	AC071669	AC071669 Giardia i
114	20	1.0	20	6	AX111291	AX111291 Sequence	c 187	19	1.0	864	33	AC039883	AC039883 Giardia i
115	20	1.0	20	6	E14593	E14593 PCR primer	c 188	19	1.0	914	33	AC089821	AC089821 Giardia i
116	20	1.0	20	6	E14594	E14594 PCR primer	c 189	19	1.0	915	33	AC046483	AC046483 Giardia i
117	20	1.0	25	6	AX110485	AX110485 Sequence	c 190	19	1.0	935	33	AC059516	AC059516 Giardia i
118	20	1.0	550	6	AX111543	AX111543 Sequence	c 191	19	1.0	1050	33	HEA28DKMP	M59804 Haemophilus
119	20	1.0	554	6	AX111326	AX111326 Sequence	192	19	1.0	1065	1	AC039210	AC039210 Giardia i
120	20	1.0	680	6	A97193	A97193 Sequence 39	193	19	1.0	1065	33	FSAAP0B	M15892 Petromyzon
121	20	1.0	1290	9	BC010387	BC010387 Homo sapi	c 194	19	1.0	1126	5	AF201456	AF201456 Arabidops
122	20	1.0	1620	14	AF425763	AF425763 Human cal	c 195	19	1.0	1922	8	AR020783	AR020783 Sequence
123	20	1.0	1620	14	AF425764	AF425764 Human cal	196	19	1.0	2193	6	AR020784	AR020784 Sequence
124	20	1.0	1620	14	AF425766	AF425766 Human cal	197	19	1.0	2193	6	AF281970	AF281970 Homo sapi
125	20	1.0	1620	14	AF427115	AF427115 Human cal	c 198	19	1.0	2195	9	AF101780	AF101780 Streptoco
126	20	1.0	1620	14	AF427116	AF427116 Human cal	199	19	1.0	2196	1	MM085259	MM085259 Mus muscu
127	20	1.0	1620	14	AF427121	AF427121 Human cal	200	19	1.0	2203	10	SPAJ2292	SPAJ2292 Streptoco
128	20	1.0	1909	5	AB024280	AB024280 Pelodiscu	c 201	19	1.0	2208	1	AX344602	AX344602 Sequence
129	20	1.0	2184	1	SMI295852	AJ295852 Streptococ	202	19	1.0	6906	6	AF288082	AF288082 Photobab
130	20	1.0	3472	10	MUSNEDD1	D10712 Mus muscu	203	19	1.0	7356	1	AF32699	AF32699 Haemophilus
131	20	1.0	40250	9	AF040250	AF040250 Homo sapi	204	19	1.0	10146	1	U32699	U32699 Streptoco
132	20	1.0	4508	9	HSPOGAGNA	Z22642 H.sapiens P	c 205	19	1.0	10856	1	AE008546	AE008546 Streptoco
133	20	1.0	4508	9	HUMRFC	L14922 Homo sapien	206	19	1.0	11486	1	U32744	U32744 Haemophilus
134	20	1.0	4952	9	HUMRFC	L23320 Human repli	c 207	19	1.0	11486	1	U32744	U32744 Haemophilus
135	20	1.0	5124	9	AE0028958	AB028958 Homo sapi	208	19	1.0	18662	3	AF381613	AF381613 Ancylost
136	20	1.0	10061	1	AE004314	AE004314 Vibrio ch	c 209	19	1.0	35447	2	AC007181	AC007181 Homo sapi
137	20	1.0	29101	3	CE2K1058	Z35604 Caenorhabdi	c 210	19	1.0	38583	2	AC017825	AC017825 Drosophil
138	20	1.0	39816	2	AC110404	AC110404 Rattus no	c 211	19	1.0	38868	3	U28941	U28941 Caenorhabdi
139	20	1.0	46555	3	CEY57G11A	Z99279 Caenorhabdi	212	19	1.0	43361	2	AC094627	AC094627 Rattus no
140	20	1.0	48060	9	AL139798	AL139798 Human DNA	213	19	1.0	45980	2	U82206	U82206 Homo sapien
141	20	1.0	57233	2	AC100137	AC100137 Mus muscu	214	19	1.0	58442	2	AC109278	AC109278 Mus muscu
142	20	1.0	66924	8	NCB9J10	AL356324 Neurospor	215	19	1.0	60543	2	AC100307	AC100307 Mus muscu
143	20	1.0	70564	9	HSBJ346G2	AL049546 Human DNA	216	19	1.0	62355	3	AC084472	AC084472 Caenorhab
144	20	1.0	84023	9	AL451131	AL451131 Human DNA	217	19	1.0	67698	2	AC105979	AC105979 Mus muscu
145	20	1.0	92348	2	AC094515	AC094515 Rattus no	c 218	19	1.0	68667	2	AL101811	AL101811 Mus muscu
146	20	1.0	108885	9	AC023135	AC023135 Homo sapi	c 219	19	1.0	73454	9	AL136988	AL136988 Human DNA

c 220	1.0	85097	3	AC026301	AC026301	Caenorhab	293	1.0	183775	2	AC093772	AC093772	Homo sapi
c 221	1.0	86950	8	AL042118	AL042118	Arabidops	c 294	1.0	185272	2	AC090734	AC090734	Homo sapi
c 222	1.0	89064	2	AL645826	AL645826	Danio rer	c 295	1.0	185528	2	AC020803	AC020803	Mus muscu
c 223	1.0	93818	2	AL627423	AL627423	Danio rer	c 296	1.0	185785	3	AC091221	AC091221	Drosophil
c 224	1.0	94236	2	AC020539	AC020539	Drosophil	c 297	1.0	187276	2	AC104247	AC104247	Homo sapi
c 225	1.0	101614	2	AC008603	AC008603	Homo sapi	c 298	1.0	188105	9	AC073655	AC073655	Homo sapi
c 226	1.0	110000	2	AC092450_3	Continuation (4 of		c 299	1.0	189036	2	AC019142	AC019142	Homo sapi
c 227	1.0	110000	2	AC026673_0	Continuation (4 of		300	1.0	190039	2	AC087639	AC087639	Homo sapi
c 228	1.0	110785	9	AC074183	AC074183	Homo sapi	301	1.0	190197	2	AC026381	AC026381	Mus muscu
c 229	1.0	117157	8	AC034258	AC034258	Oryza sat	302	1.0	191162	2	AC096629	AC096629	Bos tauru
c 230	1.0	118873	9	HS1184F4	AL034550	Human DNA	c 303	1.0	191601	2	AC092143	AC092143	Homo sapi
c 231	1.0	123530	2	AL356369	AL356369	Homo sapi	c 304	1.0	192021	2	AP003536	AP003536	Oryza sat
c 232	1.0	124280	9	HS03393J4	AL078595	Human DNA	305	1.0	194350	10	AC079832	AC079832	Mus muscu
c 233	1.0	124758	2	AC008464	AC008464	Homo sapi	306	1.0	197646	2	AC011041	AC011041	Homo sapi
c 234	1.0	125231	9	AC027296	AC027296	Homo sapi	c 307	1.0	198259	9	AC016825	AC016825	Homo sapi
c 235	1.0	126126	2	AC093304	AC093304	Homo sapi	c 308	1.0	198616	2	AC093139	AC093139	Pan trogl
c 236	1.0	131928	9	AC091738	AC091738	Homo sapi	c 309	1.0	199008	2	AL645668	AL645668	Mus muscu
c 237	1.0	132887	9	AL160408	AL160408	Human DNA	310	1.0	205050	1	AT414155	AT414155	Yersinia
c 238	1.0	133350	9	AC022511	AC022511	Homo sapi	c 311	1.0	206117	2	AC079845	AC079845	Mus muscu
c 239	1.0	133974	2	AP004359	AP004359	Oryza sat	312	1.0	207897	2	AL590389	AL590389	Mus muscu
c 240	1.0	138373	2	AC011380	AC011380	Homo sapi	c 313	1.0	212029	2	AC011318	AC011318	Homo sapi
c 241	1.0	141293	2	AP004337	AP004337	Oryza sat	c 314	1.0	225943	2	AL645934	AL645934	Mus muscu
c 242	1.0	142348	9	AP003788	AP003788	Homo sapi	c 315	1.0	237405	2	AC068142	AC068142	Mus muscu
c 243	1.0	148336	9	AC007110	AC007110	Homo sapi	316	1.0	240968	2	AC021709	AC021709	Mus muscu
c 244	1.0	150263	9	AC073218	AC073218	Homo sapi	c 317	1.0	247696	2	AC007493	AC007493	Homo sapi
c 245	1.0	151008	2	AP004261	AP004261	Oryza sat	c 318	1.0	251039	2	AL592290	AL592290	Homo sapi
c 246	1.0	152185	2	AC012219	AC012219	Homo sapi	319	1.0	252360	2	AC091272	AC091272	Mus muscu
c 247	1.0	152794	9	AL139376	AL139376	Human DNA	320	1.0	253176	3	AE003718	AE003718	Drosophil
c 248	1.0	153087	2	AC017072	AC017072	Homo sapi	321	1.0	280464	2	AC021061	AC021061	Mus muscu
c 249	1.0	153704	9	AC008567	AC008567	Homo sapi	322	1.0	298216	2	AC006875	AC006875	Caenorhab
c 250	1.0	155086	2	AC109913	AC109913	Bos tauru	323	1.0	331326	2	AC006802	AC006802	Caenorhab
c 251	1.0	155276	2	AC091992	AC091992	Homo sapi	324	1.0	339650	1	AP003583	AP003583	Nostoc sp
c 252	1.0	155334	2	AC015822	AC015822	Homo sapi	325	1.0	346296	8	AF090447	AF090447	2ea may
c 253	1.0	155827	2	AC104172	AC104172	Homo sapi	326	1.0	346296	8	AF090447	AF090447	2ea may
c 254	1.0	156077	9	AL354997	AL354997	Human DNA	327	1.0	346296	8	AF090447	AF090447	2ea may
c 255	1.0	159185	2	AC074218	AC074218	Mus muscu	c 328	1.0	346296	8	AF090447	AF090447	2ea may
c 256	1.0	159499	2	AL645612	AL645612	Homo sapi	329	1.0	346296	8	AF090447	AF090447	2ea may
c 257	1.0	159664	2	AL592441	AL592441	Homo sapi	c 330	1.0	346296	8	AF090447	AF090447	2ea may
c 258	1.0	160824	9	AC079228	AC079228	Homo sapi	c 331	1.0	346296	8	AF090447	AF090447	2ea may
c 259	1.0	162036	2	AC015773	AC015773	Homo sapi	332	1.0	346296	8	AF090447	AF090447	2ea may
c 260	1.0	163469	9	AC080009	AC080009	Homo sapi	333	1.0	346296	8	AF090447	AF090447	2ea may
c 261	1.0	163803	9	AC093531	AC093531	Homo sapi	334	1.0	346296	8	AF090447	AF090447	2ea may
c 262	1.0	164721	2	AC013657	AC013657	Homo sapi	335	1.0	346296	8	AF090447	AF090447	2ea may
c 263	1.0	164798	2	AC098686	AC098686	Bos tauru	336	1.0	346296	8	AF090447	AF090447	2ea may
c 264	1.0	165067	9	AL138927	AL138927	Human DNA	337	1.0	346296	8	AF090447	AF090447	2ea may
c 265	1.0	167075	2	AC092858	AC092858	Bos tauru	338	1.0	346296	8	AF090447	AF090447	2ea may
c 266	1.0	167642	2	AC027235	AC027235	Homo sapi	c 339	1.0	346296	8	AF090447	AF090447	2ea may
c 267	1.0	169660	9	AY050668	AY050668	Homo sapi	340	1.0	346296	8	AF090447	AF090447	2ea may
c 268	1.0	169660	9	AC011479	AC011479	Homo sapi	341	1.0	346296	8	AF090447	AF090447	2ea may
c 269	1.0	170440	2	AC048375	AC048375	Homo sapi	c 342	1.0	346296	8	AF090447	AF090447	2ea may
c 270	1.0	170944	2	AC094217	AC094217	Rattus no	343	1.0	346296	8	AF090447	AF090447	2ea may
c 271	1.0	171210	2	AC087543	AC087543	Oryza sat	c 344	1.0	346296	8	AF090447	AF090447	2ea may
c 272	1.0	171210	2	AC025905	AC025905	Oryza sat	c 345	1.0	346296	8	AF090447	AF090447	2ea may
c 273	1.0	171617	2	AC016053	AC016053	Homo sapi	c 346	1.0	346296	8	AF090447	AF090447	2ea may
c 274	1.0	172393	2	AC092873	AC092873	Pan trogl	c 347	1.0	346296	8	AF090447	AF090447	2ea may
c 275	1.0	172983	9	AC002454	AC002454	Human BAC	c 348	1.0	346296	8	AF090447	AF090447	2ea may
c 276	1.0	173807	2	AC107383	AC107383	Homo sapi	c 349	1.0	346296	8	AF090447	AF090447	2ea may
c 277	1.0	173997	2	AC012340	AC012340	Homo sapi	c 350	1.0	346296	8	AF090447	AF090447	2ea may
c 278	1.0	174296	2	AC013805	AC013805	Homo sapi	c 351	1.0	346296	8	AF090447	AF090447	2ea may
c 279	1.0	174415	2	AC104666	AC104666	Homo sapi	c 352	1.0	346296	8	AF090447	AF090447	2ea may
c 280	1.0	175440	3	AC007804	AC007804	Drosophil	c 353	1.0	346296	8	AF090447	AF090447	2ea may
c 281	1.0	175789	2	AP000425	AP000425	Homo sapi	c 354	1.0	346296	8	AF090447	AF090447	2ea may
c 282	1.0	176218	2	AC107424	AC107424	Homo sapi	c 355	1.0	346296	8	AF090447	AF090447	2ea may
c 283	1.0	179007	2	AC012275	AC012275	Homo sapi	c 356	1.0	346296	8	AF090447	AF090447	2ea may
c 284	1.0	179057	2	AC108677	AC108677	Homo sapi	c 357	1.0	346296	8	AF090447	AF090447	2ea may
c 285	1.0	179132	10	AL358892	AL358892	Mouse DNA	c 358	1.0	346296	8	AF090447	AF090447	2ea may
c 286	1.0	179364	2	AC105447	AC105447	Homo sapi	c 359	1.0	346296	8	AF090447	AF090447	2ea may
c 287	1.0	180485	2	AC027145	AC027145	Homo sapi	c 360	1.0	346296	8	AF090447	AF090447	2ea may
c 288	1.0	180672	9	AC011451	AC011451	Homo sapi	c 361	1.0	346296	8	AF090447	AF090447	2ea may
c 289	1.0	181192	2	AC106085	AC106085	Rattus no	c 362	1.0	346296	8	AF090447	AF090447	2ea may
c 290	1.0	181395	9	AC083800	AC083800	Homo sapi	c 363	1.0	346296	8	AF090447	AF090447	2ea may
c 291	1.0	182080	2	AC079682	AC079682	Mus muscu	c 364	1.0	346296	8	AF090447	AF090447	2ea may
c 292	1.0	182435	9	AC096591	AC096591	Homo sapi	c 365	1.0	346296	8	AF090447	AF090447	2ea may

c 366	18	0.9	1643	8	D38018	D38018 Beta webbia	439	18	0.9	45005	9	AC004511	AC004511 Homo sapi
c 367	18	0.9	1697	8	CRSAH11	Z26881 C.roseus SA	440	18	0.9	46509	2	AC013091	AC013091 Drosophill
c 368	18	0.9	1975	8	AK022414	AK022414 Homo sapi	c 441	18	0.9	49347	2	AC014413	AC014413 Drosophill
c 369	18	0.9	1980	8	SCU22449	U22449 Saccharomyc	c 442	18	0.9	49999	6	AX015907	AX015907 Sequence
c 370	18	0.9	2040	8	D38016	D38016 Sugar beet	c 443	18	0.9	50323	3	AE003118	AE003118 Drosophill
c 371	18	0.9	2048	1	AF274062	AF274062 Lactobaci	444	18	0.9	50974	9	AP000958	AP000958 Homo sapi
c 372	18	0.9	2068	8	F202550S01	AF202550 Homo sapi	445	18	0.9	51211	2	AC010728	AC010728 Mus muscu
c 373	18	0.9	2073	8	BVRFLPS	X87637 B.vulgaris	446	18	0.9	52149	2	AC015590	AC015590 Homo sapi
c 374	18	0.9	2074	8	BVRFLPF	X87638 B.vulgaris	447	18	0.9	52520	9	AL591592	AL591592 Human DNA
c 375	18	0.9	2083	4	OCU35428	U35428 Oryctolagus	c 448	18	0.9	54039	9	AL606803	AL606803 Human DNA
c 376	18	0.9	2089	1	ECENO	X82400 E.coli eno	c 449	18	0.9	57906	2	AC108909	AC108909 Mus muscu
c 377	18	0.9	2097	9	BC001309	BC001309 Homo sapi	c 450	18	0.9	59289	2	AC108629	AC108629 Rattus no
c 378	18	0.9	2121	8	SCYKL074C	Z28074 S.cerevisia	c 451	18	0.9	59429	2	AC091176	AC091176 Homo sapi
c 379	18	0.9	2215	3	AY069334	AY069334 Drosophill	c 452	18	0.9	60019	3	AC005650	AC005650 Drosophill
c 380	18	0.9	2293	8	NPRBCS8B	X13711 Nicotiana p	c 453	18	0.9	60379	2	AC096619	AC096619 Rattus no
c 381	18	0.9	2293	8	TOBRBCS8B	M36685 N.plumbagin	c 454	18	0.9	60761	2	AC102704	AC102704 Mus muscu
c 382	18	0.9	2391	1	BACADESYN	M83690 Bacillus su	c 455	18	0.9	62916	8	AB046427	AB046427 Arabidops
c 383	18	0.9	2500	14	D83410	D83410 Barley mild	456	18	0.9	63853	2	AC087663	AC087663 Homo sapi
c 384	18	0.9	2687	10	AB037362S1	AB037362 Mus muscu	c 457	18	0.9	63966	9	AL451141	AL451141 Human DNA
c 385	18	0.9	2700	8	AF106950	AF106950 Filobasid	c 458	18	0.9	64252	2	AC073032	AC073032 Homo sapi
c 386	18	0.9	2958	10	RNITI14	Y11283 R.norvegicu	c 459	18	0.9	64965	9	AC006962	AC006962 Homo sapi
c 387	18	0.9	3275	8	D38019	D38019 Sugar beet	460	18	0.9	66542	9	AF134576	AF134576 Homo sapi
c 388	18	0.9	3464	14	BMNAL	D10947 Barley mild	461	18	0.9	66673	9	AL138847	AL138847 Human DNA
c 389	18	0.9	3485	6	E06133	E06133 cDNA sequen	462	18	0.9	67292	2	AC102708	AC102708 Mus muscu
c 390	18	0.9	4350	8	YSCRNASE	L06209 Saccharomyc	463	18	0.9	67623	2	AC027436	AC027436 Homo sapi
c 391	18	0.9	4371	1	AF390867	AF390867 Paracoccu	c 464	18	0.9	67623	2	AC027436	AC027436 Homo sapi
c 392	18	0.9	5142	8	MIATPAI2	X15982 Ascobolus i	c 465	18	0.9	68217	9	AL590664	AL590664 Human DNA
c 393	18	0.9	5578	3	AF162798	AF162798 Drosophill	c 466	18	0.9	68792	2	AC107724	AC107724 Mus muscu
c 394	18	0.9	5941	14	AF036333	AF036333 Diatraea	c 467	18	0.9	70714	2	AC101499	AC101499 Mus muscu
c 395	18	0.9	6887	6	AX344607	AX344607 Sequence	c 468	18	0.9	71959	2	AC108436	AC108436 Mus muscu
c 396	18	0.9	7067	6	AX286953	AX286953 Sequence	c 469	18	0.9	73038	2	AC020884_3	Continuation (4 of
c 397	18	0.9	7187	2	AC014572	AC014572 Drosophill	c 470	18	0.9	73096	2	AC091063	AC091063 Homo sapi
c 398	18	0.9	7263	14	D83408	D83408 Barley mild	c 471	18	0.9	73390	9	AC012038	AC012038 Homo sapi
c 399	18	0.9	7647	8	AB027617	AB027617 Pyrus pyr	472	18	0.9	74508	5	AL591511	AL591511 Zebrafish
c 400	18	0.9	7925	3	AB042120	AB042120 Bombyx mo	c 473	18	0.9	74629	2	AC101440	AC101440 Mus muscu
c 401	18	0.9	7975	2	AC020446	AC020446 Drosophill	474	18	0.9	75031	9	AL390716	AL390716 Human DNA
c 402	18	0.9	8011	2	AC017787	AC017787 Drosophill	475	18	0.9	76388	2	AC100200	AC100200 Mus muscu
c 403	18	0.9	9347	9	HSDI156	X15909 Homo sapien	476	18	0.9	79598	2	AC012281	AC012281 Homo sapi
c 404	18	0.9	9905	2	AC014733	AC014733 Drosophill	c 477	18	0.9	82912	8	ATT21J18	ATT21J18 Arabidops
c 405	18	0.9	10006	14	AF131870	AF131870 Simian im	c 478	18	0.9	83057	9	AL445703	AL445703 Human DNA
c 406	18	0.9	10050	1	AE004335	AE004335 Vibrio ch	c 479	18	0.9	83495	3	AC087076	AC087076 Caenorhab
c 407	18	0.9	10056	1	AE006480	AE006480 Streptoco	c 480	18	0.9	84196	8	ATT3A5	ATT3A5 Arabidops
c 408	18	0.9	10059	1	AE004369	AE004369 Vibrio ch	c 481	18	0.9	84392	2	AC013393	AC013393 Homo sapi
c 409	18	0.9	10330	1	AE006776	AE006776 Sulfolobu	c 482	18	0.9	84974	8	AC007843	AC007843 Arabidops
c 410	18	0.9	10385	1	AE000361	AE000361 Escherich	c 483	18	0.9	85186	2	AC021348	AC021348 Homo sapi
c 411	18	0.9	10707	1	AE002316	AE002316 Chlamydia	c 484	18	0.9	86075	8	NC93G11	NC93G11 Neurospor
c 412	18	0.9	11142	1	AE000805	AE000805 Methanoba	c 485	18	0.9	86569	2	AC106139	AC106139 Rattus no
c 413	18	0.9	11311	1	AE005527	AE005527 Escherich	c 486	18	0.9	87397	9	AL356778	AL356778 Human DNA
c 414	18	0.9	11361	1	AE000950	AE000950 Archaeogl	c 487	18	0.9	87884	2	AC012997	AC012997 Drosophill
c 415	18	0.9	12111	1	AE006275	AE006275 Lactococc	c 488	18	0.9	89352	9	AC006030	AC006030 Homo sapi
c 416	18	0.9	12229	1	AE000913	AE000913 Methanoba	c 489	18	0.9	89732	2	AC091371	AC091371 Rattus no
c 417	18	0.9	13234	1	ECU29580	U29580 Escherichia	c 490	18	0.9	90141	9	AC035150	AC035150 Homo sapi
c 418	18	0.9	16071	1	AE000796	AE000796 Methanoba	c 491	18	0.9	90547	9	HS265J14	HS265J14 Human DNA
c 419	18	0.9	16238	1	AE007094	AE007094 Mycobacte	c 492	18	0.9	91817	2	AC109605	AC109605 Mus muscu
c 420	18	0.9	18400	3	CER13G10	Z35602 Caenorhabdi	c 493	18	0.9	92910	9	AL590368	AL590368 Human DNA
c 421	18	0.9	18689	1	AE001753	AE001753 Thermotog	c 494	18	0.9	93153	2	AF257499	AF257499 Homo sapi
c 422	18	0.9	19317	9	AL591862	AL591862 Human DNA	c 495	18	0.9	93217	2	AC023171	AC023171 Homo sapi
c 423	18	0.9	22700	2	AC020792	AC020792 Mus muscu	c 496	18	0.9	93893	9	AL669984	AL669984 Human DNA
c 424	18	0.9	23435	2	AC014167	AC014167 Drosophill	c 497	18	0.9	95091	9	AL353594	AL353594 Human DNA
c 425	18	0.9	23864	3	AE002853	AE002853 Drosophill	c 498	18	0.9	97011	9	HS997015	HS997015 Human DNA
c 426	18	0.9	23967	1	MTCY07A7	Z95556 Mycobacteri	c 499	18	0.9	97137	9	AC004745	AC004745 Homo sapi
c 427	18	0.9	29994	3	AE003032	AE003032 Drosophill	c 500	18	0.9	97700	9	HS28010	HS28010 Human DNA
c 428	18	0.9	30352	3	U28412	U28412 Caenorhabdi	c 501	18	0.9	98334	2	AC095642	AC095642 Rattus no
c 429	18	0.9	32148	8	CEW0725	Z46660 S.cerevisia	c 502	18	0.9	99228	9	AL133376	AL133376 Human DNA
c 430	18	0.9	32158	3	CEW04G3	Z68014 Caenorhabdi	c 503	18	0.9	100000	9	AP000066	AP000066 Homo sapi
c 431	18	0.9	32799	8	SPAC227	AL133156 S.pombe c	c 504	18	0.9	101188	2	AC010606	AC010606 Homo sapi
c 432	18	0.9	33330	8	AP001304	AP001304 Arabidops	c 505	18	0.9	101903	2	AP004334	AP004334 Rattus sat
c 433	18	0.9	35564	3	CEF23H12	Z74472 Caenorhabdi	c 506	18	0.9	103157	2	AC103199	AC103199 Rattus no
c 434	18	0.9	35935	9	AC092305	AC092305 Homo sapi	c 507	18	0.9	104475	9	AC011384	AC011384 Homo sapi
c 435	18	0.9	36127	3	CEF01D5	Z81493 Caenorhabdi	c 508	18	0.9	104995	9	AC020942	AC020942 Homo sapi
c 436	18	0.9	37577	8	AL590284	AL590284 Human DNA	c 509	18	0.9	104995	9	AC026416	AC026416 Homo sapi
c 437	18	0.9	37949	8	SPBC725	AL034352 S.pombe c	c 510	18	0.9	105399	9	AL589653	AL589653 Human DNA
c 438	18	0.9	38976	3	AF063007	AF063007 Caenorhab	c 511	18	0.9	105931	2	AC005427	AC005427 Drosophill

c 512	18	0.9 107529	2	AC069085	Homo sapi	AC069085	18	c 585	18	0.9 149505	9	AC093767	Homo sapi	AC093767
c 513	18	0.9 108098	9	AC008411	Homo sapi	AC008411	18	c 586	18	0.9 149699	8	AP002486	Oryza sat	AP002486
c 514	18	0.9 108500	9	AP001624	Homo sapi	AP001624	18	c 587	18	0.9 149719	8	AC0097312	Rattus no	AC0097312
c 515	18	0.9 109786	8	F5K24	Arabidops	AF128395	18	c 588	18	0.9 149745	2	AC016130	Drosophil	AC016130
c 516	18	0.9 110000	8	AC026673_0	Homo sapi	AC026673	18	c 589	18	0.9 149916	2	AL1359540	Homo sapi	AL1359540
c 517	18	0.9 111122	8	ATF11C1	Arabidops	AL132976	18	c 590	18	0.9 150424	9	AL391811	Human DNA	AL391811
c 518	18	0.9 111758	2	AC105693	Rattus no	AC105693	18	c 591	18	0.9 150606	2	AC027376	Homo sapi	AC027376
c 519	18	0.9 113194	2	AF301236	Homo sapi	AF301236	18	c 592	18	0.9 150997	9	HS11703	Human DNA	AL200995
c 520	18	0.9 113473	9	HS1040613	Human DNA	AL121887	18	c 593	18	0.9 151131	2	AP003916	Oryza sat	AP003916
c 521	18	0.9 115532	2	AC008424	Homo sapi	AC008424	18	c 594	18	0.9 151305	2	AC072029	Homo sapi	AC072029
c 522	18	0.9 115909	9	AC078901	Homo sapi	AC078901	18	c 595	18	0.9 151611	2	AP001131	Homo sapi	AP001131
c 523	18	0.9 116261	9	AL355357	Human DNA	AL355357	18	c 596	18	0.9 151679	2	AP001135	Homo sapi	AP001135
c 524	18	0.9 118068	2	AL356798	Human DNA	AL356798	18	c 597	18	0.9 151734	2	AC021225	Homo sapi	AC021225
c 525	18	0.9 118447	2	AC087397	Homo sapi	AC087397	18	c 598	18	0.9 151795	9	HS85661	Homo sapi	AC033381
c 526	18	0.9 118507	8	ATF7J8	Arabidops	AL137189	18	c 599	18	0.9 151900	9	AC096644	Human DNA	AC096644
c 527	18	0.9 118911	2	AC059063	Homo sapi	AC059063	18	c 600	18	0.9 151938	9	HS323M4	Human DNA	AC033378
c 528	18	0.9 119063	9	AC027311	Homo sapi	AC027311	18	c 601	18	0.9 151959	2	AC021689	Homo sapi	AC021689
c 529	18	0.9 119430	8	ATF9D16	Arabidops	AL035394	18	c 602	18	0.9 151987	2	AP004339	Oryza sat	AP004339
c 530	18	0.9 121496	2	AC090244	Homo sapi	AC090244	18	c 603	18	0.9 152269	2	AL591024	Homo sapi	AL591024
c 531	18	0.9 122250	9	AC005273	Homo sapi	AC005273	18	c 604	18	0.9 152336	9	AP004714	Homo sapi	AP004714
c 532	18	0.9 123010	9	AC087321	Homo sapi	AC087321	18	c 605	18	0.9 152354	2	AC095280	Rattus no	AC095280
c 533	18	0.9 123519	30	AL161639	Human DNA	AL161639	18	c 606	18	0.9 152702	2	AC106396	Homo sapi	AC106396
c 534	18	0.9 123642	2	AC094853	Rattus no	AC094853	18	c 607	18	0.9 152805	9	AC078923	Homo sapi	AC078923
c 535	18	0.9 123805	9	AC002464	Human BAC	AC002464	18	c 608	18	0.9 152837	2	AC105190	Homo sapi	AC105190
c 536	18	0.9 123837	2	AC013661	Homo sapi	AC013661	18	c 609	18	0.9 152951	9	AC026371	Homo sapi	AC026371
c 537	18	0.9 123953	2	AP003747	Oryza sat	AP003747	18	c 610	18	0.9 152996	9	AC013719	Homo sapi	AC013719
c 538	18	0.9 124368	2	AC055703	Mus muscu	AC055703	18	c 611	18	0.9 153528	2	AC024031	Homo sapi	AC024031
c 539	18	0.9 125045	2	AC094470	Rattus no	AC094470	18	c 612	18	0.9 154011	2	AC018571	Homo sapi	AC018571
c 540	18	0.9 125685	9	AC073310	Homo sapi	AC073310	18	c 613	18	0.9 154189	2	AC073180	Homo sapi	AC073180
c 541	18	0.9 126589	9	AL512444	Human DNA	AL512444	18	c 614	18	0.9 154588	9	AL1355871	Human DNA	AL1355871
c 542	18	0.9 126619	9	AC010265	Homo sapi	AC010265	18	c 615	18	0.9 154772	9	AL138764	Human DNA	AL138764
c 543	18	0.9 126910	9	AC005856	Homo sapi	AC005856	18	c 616	18	0.9 154880	2	AC084695	Mus muscu	AC084695
c 544	18	0.9 127192	9	HS106502	Human DNA	AL035562	18	c 617	18	0.9 155191	2	AC087888	Homo sapi	AC087888
c 545	18	0.9 127577	2	AC104459	Homo sapi	AC104459	18	c 618	18	0.9 155304	2	AC024713	Homo sapi	AC024713
c 546	18	0.9 127968	8	AC022492	Genomic s	AC022492	18	c 619	18	0.9 155531	2	AL592213	Homo sapi	AL592213
c 547	18	0.9 130626	9	AP001625	Homo sapi	AP001625	18	c 620	18	0.9 155577	2	AC011565	Homo sapi	AC011565
c 548	18	0.9 130737	9	AC092360	Homo sapi	AC092360	18	c 621	18	0.9 155789	2	AL669860	Mus muscu	AL669860
c 549	18	0.9 131487	2	AP004117	Oryza sat	AP004117	18	c 622	18	0.9 155842	2	AC025650	Homo sapi	AC025650
c 550	18	0.9 131692	8	AC006423	Arabidops	AC006423	18	c 623	18	0.9 156503	9	AC011511	Homo sapi	AC011511
c 551	18	0.9 135438	9	AC026444	Homo sapi	AC026444	18	c 624	18	0.9 156741	2	AL355867	Homo sapi	AL355867
c 552	18	0.9 135752	9	AC073465	Homo sapi	AC073465	18	c 625	18	0.9 156836	2	AC068059	Homo sapi	AC068059
c 553	18	0.9 136095	2	AF161326	Homo sapi	AF161326	18	c 626	18	0.9 156870	2	AC091916	Homo sapi	AC091916
c 554	18	0.9 136328	9	AC004126	Human Chr	AC004126	18	c 627	18	0.9 157711	9	AC055738	Homo sapi	AC055738
c 555	18	0.9 136915	2	AC091028	Homo sapi	AC091028	18	c 628	18	0.9 158033	9	AC018926	Homo sapi	AC018926
c 556	18	0.9 136917	2	AC087220	Oryza sat	AC087220	18	c 629	18	0.9 158392	2	AC022048	Homo sapi	AC022048
c 557	18	0.9 137008	10	AC091618	Rattus no	AC091618	18	c 630	18	0.9 158408	9	AL391821	Human DNA	AL391821
c 558	18	0.9 137897	9	AL161913	Human DNA	AL161913	18	c 631	18	0.9 158723	8	AP002869	Oryza sat	AP002869
c 559	18	0.9 138053	9	AC091146	Homo sapi	AC091146	18	c 632	18	0.9 158952	9	AC096745	Homo sapi	AC096745
c 560	18	0.9 138473	10	AC003019	Mus muscu	AC003019	18	c 633	18	0.9 158991	2	AC087673	Homo sapi	AC087673
c 561	18	0.9 138752	9	AL161627	Human DNA	AL161627	18	c 634	18	0.9 158995	2	AL590679	Homo sapi	AL590679
c 562	18	0.9 138781	2	AP002764	Homo sapi	AP002764	18	c 635	18	0.9 159003	2	AC025255	Homo sapi	AC025255
c 563	18	0.9 139904	9	HS97D16	Human DNA	AL009179	18	c 636	18	0.9 159128	2	AC040958	Homo sapi	AC040958
c 564	18	0.9 140229	8	AP003143	Oryza sat	AP003143	18	c 637	18	0.9 159183	9	CNS05TCY	Human chr	AL355887
c 565	18	0.9 140529	2	AC011486	Homo sapi	AC011486	18	c 638	18	0.9 159387	2	AC084858	Homo sapi	AC084858
c 566	18	0.9 140554	10	AC000399	Genomic s	AC000399	18	c 639	18	0.9 159507	9	AC004985	Homo sapi	AC004985
c 567	18	0.9 140921	2	AC104800	Homo sapi	AC104800	18	c 640	18	0.9 159620	2	AP002442	Homo sapi	AP002442
c 568	18	0.9 141152	2	AP004343	Oryza sat	AP004343	18	c 641	18	0.9 159930	2	AL138757	Homo sapi	AL138757
c 569	18	0.9 141456	9	AL355300	Human DNA	AL355300	18	c 642	18	0.9 160038	9	AL138755	Human DNA	AL138755
c 570	18	0.9 142040	2	AC103268	Rattus no	AC103268	18	c 643	18	0.9 160048	10	AC002109	Genomic s	AC002109
c 571	18	0.9 142239	2	AC027635	Homo sapi	AC027635	18	c 644	18	0.9 160209	9	AL138815	Human DNA	AL138815
c 572	18	0.9 143029	2	AC016319	Homo sapi	AC016319	18	c 645	18	0.9 160210	9	AL138815	Human DNA	AL138815
c 573	18	0.9 143078	9	AL445463	Human DNA	AL445463	18	c 646	18	0.9 160666	2	AC105755	Homo sapi	AC105755
c 574	18	0.9 143302	2	AP004264	Oryza sat	AP004264	18	c 647	18	0.9 160754	9	AC006213	Homo sapi	AC006213
c 575	18	0.9 143655	2	AC010808	Homo sapi	AC010808	18	c 648	18	0.9 160855	2	AC011579	Homo sapi	AC011579
c 576	18	0.9 144723	10	AC005818	Mus muscu	AC005818	18	c 649	18	0.9 160972	3	AC007723	Drosophil	AC007723
c 577	18	0.9 145089	9	AC022418	Homo sapi	AC022418	18	c 650	18	0.9 161042	2	AC098497	Rattus no	AC098497
c 578	18	0.9 145749	2	AC091359	Rattus no	AC091359	18	c 651	18	0.9 161103	9	AL392163	Human DNA	AL392163
c 579	18	0.9 146261	2	AC097112	Oryza sat	AC097112	18	c 652	18	0.9 161245	2	AP001986	Homo sapi	AP001986
c 580	18	0.9 146889	9	AL358937	Human DNA	AL358937	18	c 653	18	0.9 161313	9	AC095508	Rattus no	AC095508
c 581	18	0.9 147109	9	AL591043	Human DNA	AL591043	18	c 654	18	0.9 161549	2	AC073174	Homo sapi	AC073174
c 582	18	0.9 147404	2	AC025759	Homo sapi	AC025759	18	c 655	18	0.9 161941	2	AL645968	Mus muscu	AL645968
c 583	18	0.9 148119	2	AC010600	Homo sapi	AC010600	18	c 656	18	0.9 161970	2	AC027706	Homo sapi	AC027706
c 584	18	0.9 148335	9	HS809F4	Human DNA	AL022400	18	c 657	18	0.9 162112	2	AC011572	Homo sapi	AC011572

c 658	18	0.9 162217	9	AC068043	AC068043 Homo sapi	c 731	18	0.9 176805	2	AC107073	AC107073 Homo sapi
c 659	18	0.9 162282	9	AC094095	AC094095 Homo sapi	c 732	18	0.9 176835	2	AC099470	AC099470 Rattus no
c 660	18	0.9 162489	9	AP001572	AP001572 Homo sapi	c 733	18	0.9 176967	2	AC022742	AC022742 Homo sapi
c 661	18	0.9 162587	9	AC092805	AC092805 Homo sapi	c 734	18	0.9 177073	2	AP000934	AP000934 Homo sapi
c 662	18	0.9 162616	9	AC099539	AC099539 Homo sapi	c 735	18	0.9 177390	2	AC096113	AC096113 Rattus no
c 663	18	0.9 162786	2	AC010783	AC010783 Homo sapi	c 736	18	0.9 178171	2	AC026961	AC026961 Homo sapi
c 664	18	0.9 162919	2	AC092088	AC092088 Canis fam	c 737	18	0.9 178282	2	AC105349	AC105349 Homo sapi
c 665	18	0.9 163110	2	AC068705	AC068705 Homo sapi	c 738	18	0.9 178405	9	AC034216	AC034216 Homo sapi
c 666	18	0.9 163426	2	AC098176	AC098176 Rattus no	c 739	18	0.9 178515	2	AC009621	AC009621 Homo sapi
c 667	18	0.9 163556	9	AL592220	AL592220 Human DNA	c 740	18	0.9 178581	2	HS512L9	AL445624 Homo sapi
c 668	18	0.9 163612	2	AC103253	AC103253 Rattus no	c 741	18	0.9 178791	3	AC008218	AC008218 Drosophill
c 669	18	0.9 163706	2	AL354927	AL354927 Homo sapi	c 742	18	0.9 178806	2	AC013557	AC013557 Homo sapi
c 670	18	0.9 163796	9	AC093725	AC093725 Homo sapi	c 743	18	0.9 178821	2	AC079640	AC079640 Mus muscu
c 671	18	0.9 163997	9	AC022731	AC022731 Homo sapi	c 744	18	0.9 178902	2	AL513525	AL513525 Homo sapi
c 672	18	0.9 164383	2	AC068023	AC068023 Homo sapi	c 745	18	0.9 178933	9	AL591479	AL591479 Human DNA
c 673	18	0.9 164950	2	AC022262	AC022262 Homo sapi	c 746	18	0.9 179199	2	AC106682	AC106682 Rattus no
c 674	18	0.9 165415	2	AC016663	AC016663 Homo sapi	c 747	18	0.9 179439	2	AC018874	AC018874 Homo sapi
c 675	18	0.9 166354	2	AC078984	AC078984 Homo sapi	c 748	18	0.9 179461	2	AC019177	AC019177 Homo sapi
c 676	18	0.9 166357	9	AC026307	AC026307 Homo sapi	c 749	18	0.9 179651	9	AC005553	AC005553 Homo sapi
c 677	18	0.9 166516	9	AC012068	AC012068 Homo sapi	c 750	18	0.9 180134	2	AL672022	AL672022 Homo sapi
c 678	18	0.9 166606	2	AC020685	AC020685 Homo sapi	c 751	18	0.9 180136	1	BAC180K	D26185 B. subtilis
c 679	18	0.9 166608	2	AC024727	AC024727 Homo sapi	c 752	18	0.9 180141	2	AC025824	AC025824 Homo sapi
c 680	18	0.9 166772	2	AC098691	AC098691 Homo sapi	c 753	18	0.9 180336	2	AC021731	AC021731 Homo sapi
c 681	18	0.9 166961	2	AC067990	AC067990 Homo sapi	c 754	18	0.9 180551	9	HUAC004158	AC004158 Homo sapi
c 682	18	0.9 166962	2	AC098180	AC098180 Rattus no	c 755	18	0.9 180990	2	AC097421	AC097421 Rattus no
c 683	18	0.9 167185	2	AC080108	AC080108 Homo sapi	c 756	18	0.9 181016	2	AC073129	AC073129 Homo sapi
c 684	18	0.9 167192	2	AC097301	AC097301 Rattus no	c 757	18	0.9 181567	9	AL354811	AL354811 Human DNA
c 685	18	0.9 167296	2	AP002507	AP002507 Homo sapi	c 758	18	0.9 181835	9	AC026124	AC026124 Homo sapi
c 686	18	0.9 167497	9	AL133479	AL133479 Human DNA	c 759	18	0.9 181994	2	AC090822	AC090822 Homo sapi
c 687	18	0.9 167650	2	HS519C2	AL449210 Homo sapi	c 760	18	0.9 182150	9	HSG278N14	AL109654 Human DNA
c 688	18	0.9 167782	2	AC010943	AC010943 Homo sapi	c 761	18	0.9 182340	2	AC024417	AC024417 Homo sapi
c 689	18	0.9 167965	2	AC095250	AC095250 Rattus no	c 762	18	0.9 182427	9	AL162390	AL162390 Human DNA
c 690	18	0.9 168265	9	AC013716	AC013716 Homo sapi	c 763	18	0.9 182770	9	AC104082	AC104082 Homo sapi
c 691	18	0.9 168837	9	AC090108	AC090108 Homo sapi	c 764	18	0.9 182839	2	AC022619	AC022619 Homo sapi
c 692	18	0.9 168991	9	AC073958	AC073958 Homo sapi	c 765	18	0.9 183060	9	AC018641	AC018641 Human Chr
c 693	18	0.9 169027	2	AC099282	AC099282 Rattus no	c 766	18	0.9 183181	8	ATCHRIV19	AL161507 Arabidops
c 694	18	0.9 169460	2	AC079068	AC079068 Homo sapi	c 767	18	0.9 183430	9	AL391827	AL391827 Human DNA
c 695	18	0.9 169476	9	AC016207	AC016207 Homo sapi	c 768	18	0.9 183515	9	AC016908	AC016908 Homo sapi
c 696	18	0.9 169566	2	AC073579	AC073579 Homo sapi	c 769	18	0.9 183903	9	AP003403	AP003403 Homo sapi
c 697	18	0.9 169936	9	CNS05TDD	AL356032 Human chr	c 770	18	0.9 183960	9	AC013445	AC013445 Homo sapi
c 698	18	0.9 170022	2	AC021559	AC021559 Homo sapi	c 771	18	0.9 184271	3	AC009256	AC009256 Drosophill
c 699	18	0.9 170090	2	AC096640	AC096640 Homo sapi	c 772	18	0.9 184873	2	AC026325	AC026325 Homo sapi
c 700	18	0.9 170479	2	AC016845	AC016845 Homo sapi	c 773	18	0.9 184925	2	AC027608	AC027608 Homo sapi
c 701	18	0.9 171216	2	AC025338	AC025338 Homo sapi	c 774	18	0.9 184938	2	AP002476	AP002476 Homo sapi
c 702	18	0.9 171621	2	AC105937	AC105937 Homo sapi	c 775	18	0.9 184999	2	AL611934	AL611934 Mus muscu
c 703	18	0.9 171962	2	AC013537	AC013537 Homo sapi	c 776	18	0.9 185134	9	AC007878	AC007878 Homo sapi
c 704	18	0.9 171985	9	AL445675	AL445675 Human DNA	c 777	18	0.9 185281	2	AC011214	AC011214 Homo sapi
c 705	18	0.9 172252	2	AC093950	AC093950 Homo sapi	c 778	18	0.9 185380	2	AC016863	AC016863 Homo sapi
c 706	18	0.9 172378	2	AC021696	AC021696 Homo sapi	c 779	18	0.9 185448	2	AC092549	AC092549 Homo sapi
c 707	18	0.9 172549	2	AC072519	AC072519 Homo sapi	c 780	18	0.9 185556	9	AC019133	AC019133 Homo sapi
c 708	18	0.9 172550	2	AC072035	AC072035 Homo sapi	c 781	18	0.9 185654	2	AC069564	AC069564 Mus muscu
c 709	18	0.9 172786	2	AC105379	AC105379 Trypanoso	c 782	18	0.9 185664	10	AC006507	AC006507 Mus muscu
c 710	18	0.9 172798	9	AC026585	AC026585 Homo sapi	c 783	18	0.9 186253	9	AC020892	AC020892 Homo sapi
c 711	18	0.9 172828	9	AC009948	AC009948 Homo sapi	c 784	18	0.9 186519	2	AC073476	AC073476 Homo sapi
c 712	18	0.9 173119	9	AL359986	AL359986 Human DNA	c 785	18	0.9 186526	2	AC108067	AC108067 Homo sapi
c 713	18	0.9 173613	3	AC007475	AC007475 Drosophill	c 786	18	0.9 186540	9	AC012478	AC012478 Homo sapi
c 714	18	0.9 173792	30	AC022216	AC022216 Homo sapi	c 787	18	0.9 186919	2	AC091544	AC091544 Homo sapi
c 715	18	0.9 173947	9	AF274856	AF274856 Homo sapi	c 788	18	0.9 187038	2	AC021179	AC021179 Homo sapi
c 716	18	0.9 174487	2	AL662924	AL662924 Homo sapi	c 789	18	0.9 187517	9	HSBA192p3	AL117340 Human DNA
c 717	18	0.9 174579	2	AC016666	AC016666 Homo sapi	c 790	18	0.9 187621	2	AC079060	AC079060 Homo sapi
c 718	18	0.9 174624	2	AL160396	AL160396 Homo sapi	c 791	18	0.9 187914	9	AC104126	AC104126 Homo sapi
c 719	18	0.9 174716	2	AC094819	AC094819 Rattus no	c 792	18	0.9 189058	9	AC015724	AC015724 Homo sapi
c 720	18	0.9 174907	2	AC025455	AC025455 Homo sapi	c 793	18	0.9 189098	9	AL512430	AL512430 Human DNA
c 721	18	0.9 175271	9	AL354718	AL354718 Human DNA	c 794	18	0.9 189229	2	AC090138	AC090138 Homo sapi
c 722	18	0.9 175451	3	AC099035	AC099035 Drosophill	c 795	18	0.9 189290	2	AC106165	AC106165 Rattus no
c 723	18	0.9 175488	2	AC013556	AC013556 Homo sapi	c 796	18	0.9 189967	2	AC079553	AC079553 Mus muscu
c 724	18	0.9 175494	9	AL590667	AL590667 Human DNA	c 797	18	0.9 190174	30	AC022739	AC022739 Homo sapi
c 725	18	0.9 175563	2	AC096341	AC096341 Rattus no	c 798	18	0.9 190221	9	AC025741	AC025741 Homo sapi
c 726	18	0.9 175782	2	AC090117	AC090117 Homo sapi	c 799	18	0.9 190418	2	AP000777	AP000777 Homo sapi
c 727	18	0.9 175782	2	AL627074	AL627074 Mus muscu	c 800	18	0.9 190766	2	AC087840	AC087840 Mus muscu
c 728	18	0.9 175967	9	AC023102	AC023102 Homo sapi	c 801	18	0.9 190793	2	AC034298	AC034298 Homo sapi
c 729	18	0.9 176097	9	AL591438	AL591438 Human DNA	c 802	18	0.9 191115	2	AC069029	AC069029 Homo sapi
c 730	18	0.9 176678	30	AC036213	AC036213 Homo sapi	c 803	18	0.9 191438	3	AC008351	AC008351 Drosophill

804	18	0.9	191583	9	AC015651	Homo sapi	877	18	0.9	240825	6	AX087869	AX087869 Sequence
805	18	0.9	191704	2	AC053529	Homo sapi	c 878	18	0.9	259109	2	AL591851	AL591851 Homo sapi
806	18	0.9	191853	2	AC084151	Homo sapi	c 879	18	0.9	259894	2	HUAC002302	HUAC002302 Homo sapi
c 807	18	0.9	192440	9	CNS01DXH	Human chr	880	18	0.9	262731	3	AE003823	AE003823 Drosophil
c 808	18	0.9	192730	2	AC026676	Homo sapi	881	18	0.9	266658	1	AP002563	AP002563 Escherich
809	18	0.9	192763	3	AC007474	Drosophil	c 882	18	0.9	295054	2	AC068887	AC068887 Homo sapi
810	18	0.9	193096	2	AC104395	Homo sapi	883	18	0.9	299751	3	AE003459	AE003459 Drosophil
c 811	18	0.9	193594	2	AL662903	Mus muscu	884	18	0.9	309207	3	AE003501	AE003501 Drosophil
812	18	0.9	193622	2	AC009659	Homo sapi	885	18	0.9	305518	2	AC006879	AC006879 Caenorhab
813	18	0.9	193925	2	AC093560	Homo sapi	886	18	0.9	340000	9	AP001675	AP001675 Homo sapi
814	18	0.9	194297	2	AL607129	Mus muscu	c 887	18	0.9	340000	9	AP001746	AP001746 Homo sapi
815	18	0.9	194520	9	AL356915	Human DNA	c 888	18	0.9	348250	1	AP003592	AP003592 Nostoc sp
816	18	0.9	194601	2	AC079949	Homo sapi	889	17	0.9	27	6	A98067	A98067 Sequence 14
c 817	18	0.9	194634	2	AC020286	Drosophil	890	17	0.9	27	6	A98074	A98074 Sequence 21
818	18	0.9	194823	2	AC087369	Homo sapi	c 891	17	0.9	27	6	AR120352	AR120352 Sequence
c 819	18	0.9	194946	2	AC027362	Homo sapi	c 892	17	0.9	35	6	E39903	E39903 High-affini
820	18	0.9	195111	2	AC096981	Rattus no	c 893	17	0.9	36	6	E39902	E39902 High-affini
c 821	18	0.9	195472	2	AC018398	Homo sapi	c 894	17	0.9	145	11	HUMSWX1127	L34349 Human chrom
822	18	0.9	195616	2	AC090344	Homo sapi	895	17	0.9	207	11	KLAJ9770	AJ229770 Kluyverom
823	18	0.9	195743	2	AC016774	Homo sapi	896	17	0.9	302	3	S78746	S78746 h-hairy pai
824	18	0.9	196349	2	AC022032	Homo sapi	c 897	17	0.9	306	5	AY046281	AY046281 Podarcis
c 825	18	0.9	196671	9	AC073479	Homo sapi	c 898	17	0.9	306	5	AY046283	AY046283 Podarcis
826	18	0.9	196677	8	ATCHRIV60	Arabidops	c 899	17	0.9	306	5	AY046284	AY046284 Podarcis
827	18	0.9	197156	2	AC090345	Homo sapi	c 900	17	0.9	306	5	AY046286	AY046286 Podarcis
828	18	0.9	197322	9	AL592144	Human DNA	901	17	0.9	321	10	MMIGD4	V00789 Mouse gene
829	18	0.9	197348	3	AC080828	Drosophil	902	17	0.9	324	11	G71383	G71383 716234231FM
c 830	18	0.9	198084	9	AF225898	Homo sapi	903	17	0.9	332	8	ABE291782	AJ291782 Atropa be
c 831	18	0.9	198241	2	AP003721	Homo sapi	904	17	0.9	358	5	AB061589	AB061589 Gallus ga
c 832	18	0.9	198523	2	AC105807	Rattus no	905	17	0.9	362	5	D63959	D63959 Gallus gall
c 833	18	0.9	199000	8	AF000396S1	Beta vulg	906	17	0.9	364	5	AB003845	AB003845 Gallus ga
834	18	0.9	199199	8	ATCHRIV59	Arabidops	907	17	0.9	367	5	AF339339	AF339339 Anas plat
c 835	18	0.9	199515	2	AC027705	Homo sapi	c 908	17	0.9	368	11	BT4428564	AJ428564 Bos tauru
c 836	18	0.9	199531	9	AC094104	Homo sapi	c 909	17	0.9	376	11	G50078	G50078 SHGC-82970
837	18	0.9	199636	2	AL645855	Mus muscu	910	17	0.9	377	5	CHKIGLA14	M15144 Chicken Ig
c 838	18	0.9	199928	2	AC092500	Homo sapi	c 911	17	0.9	393	9	HSCD30L5	AF006385 Homo sapi
c 839	18	0.9	200500	9	AC093538	Homo sapi	c 912	17	0.9	396	4	AF028145	AF028145 Dusclyon
840	18	0.9	200582	2	AC008746	Homo sapi	c 913	17	0.9	405	3	AF191165	AF191165 Dendromyr
c 841	18	0.9	200799	2	AL445648	Homo sapi	c 914	17	0.9	405	9	AF274755S5	AF274755 Homo sapi
c 842	18	0.9	200909	2	AC103915	Homo sapi	915	17	0.9	419	1	SPPBFLAB	Z21809 S.pneumonia
843	18	0.9	202929	3	AC007896	Drosophil	916	17	0.9	425	11	G53163	G53163 SHGC-84696
c 844	18	0.9	205327	2	AC107396	Homo sapi	917	17	0.9	451	3	U91383	U91383 Rivellia qu
c 845	18	0.9	205860	9	CNS01RGN	Human chr	c 918	17	0.9	468	11	G64794	G64794 B141E16 GSS
846	18	0.9	206258	2	AC073916	Homo sapi	c 919	17	0.9	494	1	L51ADPD	M80347 Listeria in
c 847	18	0.9	207945	2	CNS01DSS	Human chr	c 920	17	0.9	495	10	MMU296878	AJ296878 Mus muscu
848	18	0.9	208729	2	AC012052	Homo sapi	921	17	0.9	498	8	PDY15429	Y15429 Populus del
c 849	18	0.9	210350	2	AC079536	Mus muscu	c 922	17	0.9	511	6	AX332740	AX332740 Sequence
c 850	18	0.9	210400	2	AC108141	Homo sapi	c 923	17	0.9	511	6	AX332912	AX332912 Sequence
c 851	18	0.9	211127	2	AC067856	Homo sapi	924	17	0.9	537	5	LP1249753	AJ249753 Lygodacty
c 852	18	0.9	211628	2	AC099425	Rattus no	c 925	17	0.9	560	1	BCE243161	AJ243161 Bacillus
853	18	0.9	212360	2	AL606908	Mus muscu	c 926	17	0.9	560	1	BCE243162	AJ243162 Bacillus
854	18	0.9	213447	2	AL354875	Homo sapi	c 927	17	0.9	560	1	BCE243163	AJ243163 Bacillus
855	18	0.9	214192	2	AL646044	Mus muscu	c 928	17	0.9	560	1	BTH243147	AJ243147 Bacillus
856	18	0.9	215177	2	AC011667	Homo sapi	c 929	17	0.9	560	1	BTH243149	AJ243149 Bacillus
c 857	18	0.9	215534	1	BSUB0021	Bacillus su	930	17	0.9	560	6	AX111540	AX111540 Sequence
c 858	18	0.9	215717	2	AC092196	Canis fam	931	17	0.9	560	6	AX111542	AX111542 Sequence
c 859	18	0.9	218233	3	AC099669	Homo sapi	c 932	17	0.9	564	14	HIM404104	AJ404104 Human imm
860	18	0.9	218286	2	AC099727	Mus muscu	933	17	0.9	601	5	AF339325	AF339325 Pelaeance
861	18	0.9	218807	2	AC009175	Homo sapi	934	17	0.9	602	5	AF339324	AF339324 Pelaeance
862	18	0.9	218941	2	AC092950	Homo sapi	935	17	0.9	603	1	AB036433	AB036433 Unculture
863	18	0.9	219180	2	AC092251	Mus muscu	c 936	17	0.9	616	8	AF231284	AF231284 Leucoloma
c 864	18	0.9	220260	2	AC103535	Rattus no	c 937	17	0.9	616	8	AF231285	AF231285 Leucoloma
c 865	18	0.9	220699	2	AC009206	Drosophil	c 938	17	0.9	616	8	AF231286	AF231286 Leucoloma
c 866	18	0.9	221276	2	AC098721	Mus muscu	c 939	17	0.9	646	8	AF382710	AF382710 Thibaudia
867	18	0.9	222235	3	AC074148	Mus muscu	940	17	0.9	648	5	AF168056	AF168056 Opisthoco
868	18	0.9	224100	3	AE003770	Drosophil	941	17	0.9	657	10	MUSIGCD16	J00449 Mouse germi
c 869	18	0.9	226502	2	AL645802	Mus muscu	942	17	0.9	668	10	MUSIGKAG3	L00031 Mouse Ig ka
870	18	0.9	226916	9	AC005066	Homo sapi	943	17	0.9	706	6	AX154845	AX154845 Sequence
c 871	18	0.9	228283	10	AF312994	Mus muscu	944	17	0.9	718	3	AF146717	AF146717 Dorymyrme
c 872	18	0.9	229122	3	AE003700	Drosophil	945	17	0.9	722	5	GDSCABL1	Z46722 G.domesticu
c 873	18	0.9	229380	2	AC079636	Mus muscu	946	17	0.9	724	33	AC061416	AC061416 Gardia i
874	18	0.9	229518	2	AC069058	Homo sapi	c 947	17	0.9	728	8	TOMMT206	DB4426 Tomato mito
c 875	18	0.9	229518	2	AC069058	Homo sapi	948	17	0.9	730	8	PHZ78451	Z78451 P.hookerae
876	18	0.9	239171	3	AE003765	Drosophil	c 949	17	0.9	752	11	G64645	G64645 B141A22/T7

c 950	17	0.9	783	6	A01796	A01796 Synthetic g
c 951	17	0.9	788	9	HSN801164	AL117628 Homo sapi
c 952	17	0.9	814	33	AC062780	Ac062780 Giardia i
c 953	17	0.9	820	33	AC032294	Ac032294 Giardia i
c 954	17	0.9	834	6	AX106224	AX106224 Sequence
c 955	17	0.9	834	6	AX140515	AX140515 Sequence
c 956	17	0.9	834	6	AX200375	AX200375 Sequence
c 957	17	0.9	834	6	AX267031	AX267031 Sequence
c 958	17	0.9	834	33	AC053681	Ac053681 Giardia i
c 959	17	0.9	881	33	AC034903	Ac034903 Giardia i
c 960	17	0.9	888	33	AC064268	Ac064268 Giardia i
c 961	17	0.9	913	33	AC086182	Ac086182 Giardia i
c 962	17	0.9	921	33	AC081013	Ac081013 Giardia i
c 963	17	0.9	925	33	AC042131	Ac042131 Giardia i
c 964	17	0.9	931	33	AC082542	Ac082542 Giardia i
c 965	17	0.9	965	33	AC033632	Ac033632 Giardia i
c 966	17	0.9	968	33	AC067501	Ac067501 Giardia i
c 967	17	0.9	971	6	AR175273	AR175273 Sequence
c 968	17	0.9	972	6	AX349609	AX349609 Sequence
c 969	17	0.9	972	8	ANMP	X52565 A. nidulans
c 970	17	0.9	988	33	AC066160	Ac066160 Giardia i
c 971	17	0.9	999	33	AC065913	Ac065913 Giardia i
c 972	17	0.9	1000	8	TROPRTG	X55879 T. reesei u
c 973	17	0.9	1001	33	AC054213	Ac054213 Giardia i
c 974	17	0.9	1003	5	AF125514	AF125514 Ocorrhync
c 975	17	0.9	1003	5	AF125515	AF125515 Ocorrhync
c 976	17	0.9	1005	14	SCL290434	AJ290434 Subterranean
c 977	17	0.9	1009	8	U01020	U01020 Medicago sa
c 978	17	0.9	1012	33	AC086876	Ac086876 Giardia i
c 979	17	0.9	1031	33	AC043181	Ac043181 Giardia i
c 980	17	0.9	1032	8	MIORF206	X74164 O. berterian
c 981	17	0.9	1036	5	AF044132	AF044132 Ocorrhync
c 982	17	0.9	1036	5	AF044154	AF044154 Ocorrhync
c 983	17	0.9	1037	5	AF044131	AF044131 Ocorrhync
c 984	17	0.9	1037	5	AF044135	AF044135 Ocorrhync
c 985	17	0.9	1037	5	AF044140	AF044140 Ocorrhync
c 986	17	0.9	1037	5	AF044142	AF044142 Ocorrhync
c 987	17	0.9	1037	5	AF044143	AF044143 Ocorrhync
c 988	17	0.9	1037	5	AF044144	AF044144 Ocorrhync
c 989	17	0.9	1037	5	AF044151	AF044151 Ocorrhync
c 990	17	0.9	1037	5	AF044152	AF044152 Ocorrhync
c 991	17	0.9	1037	5	AF044153	AF044153 Ocorrhync
c 992	17	0.9	1037	5	AF044156	AF044156 Ocorrhync
c 993	17	0.9	1037	5	AF044162	AF044162 Ocorrhync
c 994	17	0.9	1037	5	AF044163	AF044163 Ocorrhync
c 995	17	0.9	1037	5	AF044164	AF044164 Ocorrhync
c 996	17	0.9	1038	5	AF044130	AF044130 Ocorrhync
c 997	17	0.9	1038	5	AF044133	AF044133 Ocorrhync
c 998	17	0.9	1038	5	AF044134	AF044134 Ocorrhync
c 999	17	0.9	1038	5	AF044139	AF044139 Ocorrhync
1000	17	0.9	1038	5	AF044148	AF044148 Ocorrhync

ALIGNMENTS

RESULT	1					
LOCUS	AR120238	1999 bp	DNA	linear	PAT 16-MAY-2001	
DEFINITION	Sequence 1 from patent US 6159469.					
ACCESSION	AR120238					
VERSION	AR120238.1	GI:14103814				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	Choi, G. H., Kunsch, C. A., Barash, S. C., Dillon, P. J., Dougherty, B.,					
AUTHORS	Fannon, M. R. and Rosen, C. A.					
TITLE	Streptococcus pneumoniae antigens and vaccines					
JOURNAL	Patent: US 6159469-A 1 12-DEC-2000;					
FEATURES	Location/Qualifiers					
source	1. .1999					

BASE COUNT	656 a	485 c	376 g	482 t	
ORIGIN	/organism="unknown"				
Query Match	100.0%; Score 1999; DB 6; Length 1999;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 1999; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Oy	1	TAAATCTACGACAATAAAATCAATCTGCTGACCTGGGTCTTGAACGCGCGTCAA	60		
Db	1	TAAATCTACGACAATAAAATCAATCTGCTGACCTGGGTCTTGAACGCGCGTCAA	60		
Oy	61	TGCCCAAGCTATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	120		
Db	61	TGCCCAAGCTATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	120		
Oy	121	TGGCTTCTTCGACACAGGGGATGTATACATCGTATCTCTGGGAGCTTCTTGGGCAA	180		
Db	121	TGGCTTCTTCGACACAGGGGATGTATACATCGTATCTCTGGGAGCTTCTTGGGCAA	180		
Oy	181	TCGCAAAAGCAATTCCTCCAAGTGGATCAACTCTCACCCCAACAGTTGATTAAAGTTGAC	240		
Db	181	TCGCAAAAGCAATTCCTCCAAGTGGATCAACTCTCACCCCAACAGTTGATTAAAGTTGAC	240		
Oy	241	TTACTTTTCAACTTCGACTTCGACACAGATATTTCTCGTAAGGCTCAGGAAGCTTGGTT	300		
Db	241	TTACTTTTCAACTTCGACTTCGACACAGATATTTCTCGTAAGGCTCAGGAAGCTTGGTT	300		
Oy	301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360		
Db	301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360		
Oy	361	GGTCTACATGCTAATGGGAATATGGGAATGCAAGCAGCAGCTCAAACTACTATGGTAA	420		
Db	361	GGTCTACATGCTAATGGGAATATGGGAATGCAAGCAGCAGCTCAAACTACTATGGTAA	420		
Oy	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGAATGCCCTCAGGACAC	480		
Db	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGAATGCCCTCAGGACAC	480		
Oy	481	AAACAATATGACCCCTATTACATCCAGACAGCCAGCCAGCCGCGGAACTTGGTCTT	540		
Db	481	AAACAATATGACCCCTATTACATCCAGACAGCCAGCCAGCCGCGGAACTTGGTCTT	540		
Oy	541	ATCTCAAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	600		
Db	541	ATCTCAAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	600		
Oy	601	ACCAATTTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTTACCTTGTACATGA	660		
Db	601	ACCAATTTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTTACCTTGTACATGA	660		
Oy	661	TAAATACCTCAAGGAAGTCAATCAAGTTGAAGAAGCAAGGCTATAACCTACTAC	720		
Db	661	TAAATACCTCAAGGAAGTCAATCAAGTTGAAGAAGCAAGGCTATAACCTACTAC	720		
Oy	721	TAACAATACAGACGAATACGTTGCCCTATCCAGACGATGAATTTGCAAGTGCCTTCTACCAT	780		
Db	721	TAACAATACAGACGAATACGTTGCCCTATCCAGACGATGAATTTGCAAGTGCCTTCTACCAT	780		
Oy	781	TTACCAATACAGACGAATACGTTGCCCTATCCAGACGATGAATTTGCAAGTGCCTTCTACCAT	840		
Db	781	TTACCAATACAGACGAATACGTTGCCCTATCCAGACGATGAATTTGCAAGTGCCTTCTACCAT	840		
Oy	841	TGTTGATGTTTCAACGGTAAAGTCAATTCGCCAGGTAGGAGCACGCCATCAGTCAAGTAA	900		
Db	841	TGTTGATGTTTCAACGGTAAAGTCAATTCGCCAGGTAGGAGCACGCCATCAGTCAAGTAA	900		
Oy	901	TGTTTTCCTCGAATTAACCAAGCAGTACAAACCAACCGGACTGGGATCAACTATGAA	960		
Db	901	TGTTTTCCTCGAATTAACCAAGCAGTACAAACCAACCGGACTGGGATCAACTATGAA	960		

QY 961 ACCGATCACAGACTATGCTCCTCGCTGGAGTACGGTGTCTAGCAATCAACTGCTACTAT 1020
Db 961 ACCGATCACAGACTATGCTCCTCGCTGGAGTACGGTGTCTAGCAATCAACTGCTACTAT 1020
QY 1021 CGTTACAGTACGCCCTATAACTACCTCGGCAAAATACCTCTGTTTATAACTTGGGATAG 1080
Db 1021 CGTTACAGTACGCCCTATAACTACCTCGGCAAAATACCTCTGTTTATAACTTGGGATAG 1080
QY 1081 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAAATCGCAAACTGCCAGC 1140
Db 1081 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAAATCGCAAACTGCCAGC 1140
QY 1141 CGTGGAACTCTAAACAAGTCGAGCTCAACCGCGCAAGACTTTCTTAATGGTCTAGG 1200
Db 1141 CGTGGAACTCTAAACAAGTCGAGCTCAACCGCGCAAGACTTTCTTAATGGTCTAGG 1200
QY 1201 AATCGACTACCCCAAGTATTCACTACTCAATGCGCATTTCAAGTAACACACCGAATCAGA 1260
Db 1201 AATCGACTACCCCAAGTATTCACTACTCAATGCGCATTTCAAGTAACACACCGAATCAGA 1260
QY 1261 CAAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGCTGCTTACGCTGCGCTTTGCAATGG 1320
Db 1261 CAAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGCTGCTTACGCTGCGCTTTGCAATGG 1320
QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGTCTTTAGTATGGGAGTGAATA 1380
Db 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGTCTTTAGTATGGGAGTGAATA 1380
QY 1381 AGAGTTCTTAATGTGCGAACTCGTCCCATGAAGGAACGACAGCCATATATGATGCCGA 1440
Db 1381 AGAGTTCTTAATGTGCGAACTCGTCCCATGAAGGAACGACAGCCATATATGATGCCGA 1440
QY 1441 CATGATGAACACAGCTTGTGACTTATGGAATGACCAATGCTATCTTCTGCTGGCTCCC 1500
Db 1441 CATGATGAACACAGCTTGTGACTTATGGAATGACCAATGCTATCTTCTGCTGGCTCCC 1500
QY 1501 TCAGGCTGTATAACAGGAACCTCTAATATACAGACGAGGAAATGGAACACATCAA 1560
Db 1501 TCAGGCTGTATAACAGGAACCTCTAATATACAGACGAGGAAATGGAACACATCAA 1560
QY 1561 GACCTCTCAATTTGTAGCACTGTATCAATATTTGCTGGCTATACGCGTAAATATTCAT 1620
Db 1561 GACCTCTCAATTTGTAGCACTGTATCAATATTTGCTGGCTATACGCGTAAATATTCAT 1620
QY 1621 GCGTGTATGACAGGCTATTCTAACCGCTGTACACCACTTTGAGGCAATGGCCTTACGGT 1680
Db 1621 GCGTGTATGACAGGCTATTCTAACCGCTGTACACCACTTTGAGGCAATGGCCTTACGGT 1680
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAGGAAGCAATCCAGAA 1740
Db 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAGGAAGCAATCCAGAA 1740
QY 1741 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGATTTTAAATGGTCTCG 1800
Db 1741 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGATTTTAAATGGTCTCG 1800
QY 1801 TTTCTAGTGGAACTCACCTGCTCCCAACAACCCCATCAACTGAAAGTTTCAAGCTCATC 1860
Db 1801 TTTCTAGTGGAACTCACCTGCTCCCAACAACCCCATCAACTGAAAGTTTCAAGCTCATC 1860
QY 1861 ATCAGATAGTCAACTTACAGTCTAGCTCAACCACTCCAAGCACAATAATAGTACGAC 1920
Db 1861 ATCAGATAGTCAACTTACAGTCTAGCTCAACCACTCCAAGCACAATAATAGTACGAC 1920
QY 1921 TACCAATCTTAACATAATACGCAACATCAATACACCCCTGATCAACAAATCAGAA 1980
Db 1921 TACCAATCTTAACATAATACGCAACATCAATACACCCCTGATCAACAAATCAGAA 1980
QY 1981 TCCCTCAACGACCAACCA 1999
Db 1981 TCCCTCAACGACCAACCA 1999

RESULT 2
BD003825/c
LOCUS
DEFINITION Polynucleotide of Streptococcus pneumoniae and sequence.
ACCESSION BD003825
VERSION BD003825.1 GI:18631786
KEYWORDS JP 2001501833-A/145.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 10711)
AUTHORS Kunsch,C.A., Choi,G.H., Dillon,P.J., Rosen,C.A., Bara,S.C.,
Fannon,M. and Dougherty,B.A.
TITLE Polynucleotide of Streptococcus pneumoniae and sequence
JOURNAL Patent: JP 2001501833-A 145 13-FEB-2001;
HUMAN GENOME SCIENCES INC
COMMENT OS Unidentified
PN JP 2001501833-A/145
PD 13-FEB-2001
PF 30-OCT-1997 JP 1998520718
PI CHARLES A KUNSCH,GIL H CHOI,PATRICK J DILLON,CRAIG A ROSEN, PI
STEVEN C BARASH,
PI MICHAEL FANNON,BRIAN A DOUGHERTY
PC C12N15/09,A01K67/027,C07K16/12,C12N1/15,C12N1/19,
PC C12N1/21,
PC C12N5/10,C12P21/02,C12Q1/68,G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..10711
FT /organism='Unidentified'.
FEATURES
source
1..10711
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 3280 a 2056 c 2412 g 2963 t
ORIGIN

Query Match 100.0%; Score 1999; DB 6; Length 10711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TAAATCTACGACAAATAAAATCAACTCATGTGCTGACTTGGGTTCTGAACGGCGCTCAA 60
Db 9767 TAAATCTACGACAAATAAAATCAACTCATGTGCTGACTTGGGTTCTGAACGGCGCTCAA 9708
QY 61 TGCCCAAGCTAATGATATTTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 9707 TGCCCAAGCTAATGATATTTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 9648
QY 121 TCGCTTCTTCGACCAAGGGGATTCATACCATCCCTATCTCGGGAGCTTCTTGGCGAA 180
Db 9647 TCGCTTCTTCGACCAAGGGGATTCATACCATCCCTATCTCGGGAGCTTCTTGGCGAA 9588
QY 181 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCACACAGTTGATTAAGTTGAC 240
Db 9587 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCACACAGTTGATTAAGTTGAC 9528
QY 241 TTACTTTTCAACTTCCGACCTTCGACCAAGCTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 9527 TTACTTTTCAACTTCCGACCTTCGACCAAGCTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 9468
QY 301 AGCGATTAGTGTAGACAAAAAGCAACCAAGCAAGAAATCTTGCACCTACTATATAATAA 360
Db 9467 AGCGATTAGTGTAGACAAAAAGCAACCAAGCAAGAAATCTTGCACCTACTATATAATAA 9408
QY 361 GGTCTACATGTCTAATGGGAACATATGGAATGCAGACAGCTCAAACTACTATGGTAA 420
Db 9407 GGTCTACATGTCTAATGGGAACATATGGAATGCAGACAGCTCAAACTACTATGGTAA 9348

QY 421 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCCTTGGCTGGATGGAATGCCCTCAGGCACC 480
|||||
Db 9347 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCCTTGGCTGGATGGAATGCCCTCAGGCACC 9288
|||||
QY 481 AAACCAATATGACCCCTATTACATCCAGAACAGCCGCGGAACTTGGTCTTT 540
|||||
Db 9287 AAACCAATATGACCCCTATTACATCCAGAACAGCCGCGGAACTTGGTCTTT 9228
|||||
QY 541 ATCTGAATGAAATTCAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
|||||
Db 9227 ATCTGAATGAAATTCAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 9168
|||||
QY 601 ACCAATTACTGATGAGTACAAAGTCTCAATACAGCAAGTAATACCTGCTTACATGGA 660
|||||
Db 9167 ACCAATTACTGATGAGTACAAAGTCTCAATACAGCAAGTAATACCTGCTTACATGGA 9108
|||||
QY 661 TAATTACCTCAAGGAGTCAATCAATCAAGTTGAGAGAAACAGGCTATAAAGCTACTAC 720
|||||
Db 9107 TAATTACCTCAAGGAGTCAATCAATCAAGTTGAGAGAAACAGGCTATAAAGCTACTAC 9048
|||||
QY 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
|||||
Db 9047 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 8988
|||||
QY 781 TTACATACAGACGAATAGCTTGGCTATCCAGACGATGAATTCAGAGTGCCTTACCAT 840
|||||
Db 8987 TTACATACAGACGAATAGCTTGGCTATCCAGACGATGAATTCAGAGTGCCTTACCAT 8928
|||||
QY 841 TGTGTATGTTTCTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
|||||
Db 8927 TGTGTATGTTTCTAACGGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 8868
|||||
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCGCGACTGGGGATCAACTATGAA 960
|||||
Db 8867 TGTTCCTTCGGAATTAACCAAGCAGTAGAACAACCGCGACTGGGGATCAACTATGAA 8808
|||||
QY 961 ACGATACAGACTATGCTTCCTGCTTGGAGTACGGTGTCTAGATTCAGTCTACTAT 1020
|||||
Db 8807 ACGATACAGACTATGCTTCCTGCTTGGAGTACGGTGTCTAGATTCAGTCTACTAT 8748
|||||
QY 1021 CGTTCAGATGAGCCCTATAACTACCTGGGCAATACCTCTGTTTATAACTGGGATAG 1080
|||||
Db 8747 CGTTCAGATGAGCCCTATAACTACCTGGGCAATACCTCTGTTTATAACTGGGATAG 8688
|||||
QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGCGAAAGCTCCACG 1140
|||||
Db 8687 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGCGAAAGCTCCACG 8628
|||||
QY 1141 CGTGGAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1200
|||||
Db 8627 CGTGGAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 8568
|||||
QY 1201 AATCGACTACCAAGTATTTCATCTACTCAATGCCATTTTCAAGTAAACAACCGCAATCAGA 1260
|||||
Db 8567 AATCGACTACCAAGTATTTCATCTACTCAATGCCATTTTCAAGTAAACAACCGCAATCAGA 8508
|||||
QY 1261 CAAAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTGCTACGCTGCCTTTGCAATGG 1320
|||||
Db 8507 CAAAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTGCTACGCTGCCTTTGCAATGG 8448
|||||
QY 1321 TGGAACTTACTATAACCAATGTATATCCATAAGTCGCTTTAGTGATGGAGTGAANA 1380
|||||
Db 8447 TGGAACTTACTATAACCAATGTATATCCATAAGTCGCTTTAGTGATGGAGTGAANA 8388
|||||
QY 1381 AGAGTTCTCTAATGTCGGAACTCGTGCCATGAAGAAACGACAGCCTATATGATGACCGA 1440
|||||
Db 8387 AGAGTTCTCTAATGTCGGAACTCGTGCCATGAAGAAACGACAGCCTATATGATGACCGA 8328
|||||
QY 1441 CATGATGAAACAGCTTTGACTTATGGAAGTGGAGCAATGCCTATCTTGTCTGGCTCCC 1500
|||||
Db 8327 CATGATGAAACAGCTTTGACTTATGGAAGTGGAGCAATGCCTATCTTGTCTGGCTCCC 8268
|||||
QY 1501 TCAGCTGGTAAACAGGAACTCTAATATACAGACGAGGAAATTTGAAACCAATCA 1560

Db 8267 TCAGCTGGTAAACAGGAACTCTAATATACAGACGAGGAAATTTGAAACCAATCA 8208
|||||
QY 1561 GACCTCTCAATTTGTAGCACCTGATGAACATATTGCTGGCTATACGCGTAAATATTCAAT 1620
|||||
Db 8207 GACCTCTCAATTTGTAGCACCTGATGAACATATTGCTGGCTATACGCGTAAATATTCAAT 8148
|||||
QY 1621 GGCTGTATGACGAGGCTATTCTAAACGCTGTACACACCTTTAGGCAATGGCCTTACGGT 1680
|||||
Db 8147 GGCTGTATGACGAGGCTATTCTAAACGCTGTACACACCTTTAGGCAATGGCCTTACGGT 8088
|||||
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTTACCTGCTGTGAAGGAAGCAATCCAGA 1740
|||||
Db 8087 CGCTGCCAAAGTTTACCGCTCTATGATGACCTTACCTGCTGTGAAGGAAGCAATCCAGA 8028
|||||
QY 1741 TTGGATATATACAGAGGGCTCTACAGAAATGGAGAAATTCGTAATTTAAAAATGGTCTCG 1800
|||||
Db 8027 TTGGATATATACAGAGGGCTCTACAGAAATGGAGAAATTCGTAATTTAAAAATGGTCTCG 7968
|||||
QY 1801 TTCTAGTGGAACTACCTGCTCCACAACACCCCATCAACTGAAAGTTCAAGCTCATC 1860
|||||
Db 7967 TTCTAGTGGAACTACCTGCTCCACAACACCCCATCAACTGAAAGTTCAAGCTCATC 7908
|||||
QY 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAAGCACAATAATAGTAGCAG 1920
|||||
Db 7907 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAAGCACAATAATAGTAGCAG 7848
|||||
QY 1921 TACCAATCTCAATAATACCAACAATCAATAACAACCCCTGATCAACAATAATCAGAA 1980
|||||
Db 7847 TACCAATCTCAATAATACCAACAATCAATAACAACCCCTGATCAACAATAATCAGAA 7788
|||||
QY 1981 TCCTCAACGACGACAACCA 1999
|||||
Db 7787 TCCTCAACGACGACAACCA 7769
|||||
RESULT 3
AE007349/c
LOCUS
DEFINITION Streptococcus pneumoniae TIGR4 section 32 of 194 of the complete genome. BCT 31-AUG-2001
ACCESSION AE007349 AE005672
VERSION AE007349.1 GI:14971837
KEYWORDS
SOURCE Streptococcus pneumoniae TIGR4
ORGANISM Streptococcus pneumoniae TIGR4
REFERENCE 1 (bases 1 to 11334)
AUTHORS Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Salzberg, S.L., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Utterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Dickson, F., Hickey, E.K., Holt, I.E., Loftus, B.J., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.
TITLE Complete genome sequence of a virulent isolate of Streptococcus pneumoniae
JOURNAL Science, 293 (5529), 498-506 (2001)
MEDLINE 21357209
PUBMED 11463916
REFERENCE 2 (bases 1 to 11334)
AUTHORS Tettelin, H., Nelson, K.E., Paulsen, I.T., Eisen, J.A., Read, T.D., Peterson, S., Heidelberg, J., DeBoy, R.T., Haft, D.H., Dodson, R.J., Durkin, A.S., Gwinn, M., Kolonay, J.F., Nelson, W.C., Peterson, J.D., Umayam, L.A., White, O., Lewis, M.R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A.M., Utterback, T.R., Hansen, C.L., McDonald, L.A., Feldblyum, T.V., Angiuoli, S., Gesuwan, P., Hickey, E.K., Holt, I.E., Loftus, B.J., Ujwal, M.L., Yang, F., Smith, H.O., Venter, J.C., Dougherty, B.A., Morrison, D.A., Hollingshead, S.K. and Fraser, C.M.
TITLE Direct Submission

JOURNAL	Submitted (29-JUN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES	Location/Qualifiers
source	1..11334
	/organism="Streptococcus pneumoniae TIGR4"
	/strain="TIGR4"
	/db_xref="taxon:170187"
gene	complement(114..2273)
	/gene="Sp0369"
CDS	complement(114..2273)
	/gene="Sp0369"
	/codon_start=1
	/transl_table=11
	/product="penicillin-binding protein 1A"
	/protein_id="AAK74536.1"
	/db_xref="GI:14971838"
	/translation="MNKPTILRLIKYLSIFLSILVIAAIVLGGGVFFYVSKAPSLSE SKLVATSSKIYDNKNQIADLGSERRVNAQNDIPTDLVKAIVSIEDHRRFDRGID TIRILGAFRLNQSLOGSTLTQQLIKLITFSTSDQTSIRKAQAWLAIQLEQK ATKQELIYYINKVYMSNGYGMQTAQNYGKDLNNLSLPQLALLAGMPAPNQDP YSHPAAQDRNLVLESENKQGYISAEQYKAVNTPITDGLQSLKASNYPAYMDNYL KEVINQVBEETGYNLLTTGMDVYTNVQEAQKHLWDIYNTDEYVAYPDELQVASTIV DVSKGVYIAQLGARHOSNVSEFIGNQAVETNRDMGSTMKPIITDVAPELEYGYDSTAT IVHDEPNVPCNTPNVYNKDRGYEGNITFLOYALQOSRVNPAVETLNKVLNKAETFLN GLGIDYPSIHYSNAISSNTESDKKYGASSEKMAAAYAAFAANGTYIRPMTIHKVYFS DGSKEFNSVGTWRAMKETATYMMTDMKTVITGYGRNAYLAWLPQAGKTGTSTNDE BIENHIKTSQVAPDELFAGYTRYKSMAVVTGYSNRLTPLVNGLTVAAKYRSNMTY LSEGSNPDNMNIPGLYRNGEFVKNGARSTWNSPAPQPPSTESSSSSSSSTSSQSS STPTSTNNSTTTNNNTTQQSNTTPOQONQNPQAP"
gene	complement(2270..2866)
	/gene="Sp0370"
CDS	complement(2270..2866)
	/gene="Sp0370"
	/codon_start=1
	/transl_table=11
	/product="recombination protein U"
	/protein_id="AAK74537.1"
	/db_xref="GI:14971839"
	/translation="MVNYPHKVSSQKROTSISQPKNFANRGMSEKMNATNDYYLSQ GLAVTHKKPTQIQVVDYQPSRAKQIVAEYFRQASTDYSGVNGYVIDFEVKETKO KRAIPMKNFHQIQHMEQVLAQOCICFVLLHFSSQETYLPLPDLIRFVHQDGQK SMPLEYIREGYEIKAGAFPOIPVLNVIKEHLGGKTR"
gene	2932..3459
	/gene="Sp0371"
CDS	2932..3459
	/gene="Sp0371"
	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="conserved hypothetical protein"
	/protein_id="AAK74538.1"
	/db_xref="GI:14971840"
	/translation="MATVALVIGYSAFDLGFSDKDPRLKLIKAIKIRKOLEMAADGVS WLVTGSLGFEHWVLEVAQEMKTEYGQATIFAFETHGENNENGNOMKLSRFQVDF VKYAYPRYEHGQLRDYQQFLLENTTSYLYFDENETKLAYFYQKKMKNQDYFKRL TFDELNEAENFSEN"
gene	3529..3858
	/gene="Sp0372"
CDS	3529..3858
	/gene="Sp0372"
	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="conserved hypothetical protein"
	/protein_id="AAK74539.1"
	/db_xref="GI:14971841"
	/translation="WASLIIFSAKDIFEQEFGREVRGYNKVEVDFELDVIKDYETAA LVKSLRQIADLKBELTRKPPSPVQAEPLAAITSSMTNFDILRLRLRLEKEVFGKO ILDNSDF"
gene	4344..5501
	/gene="Sp0373"
CDS	4344..5501
	/gene="Sp0373"
	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11

	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="conserved hypothetical protein"
	/protein_id="AAK74540.1"
	/db_xref="GI:14971842"
	/translation="MKKEFNLIATVAAGLEAVVGREVRELGYDQOVENGVRVRFQGDVR AITETNLWLRADRIKIIVGTFPAKTFEELFOGVFALDWENYLPILGARFPISKACVK SKLHNPSVQAISSKAVIKKQHYARPEGVPLMGEPFLKIEVLSILKDVAFVIMDTT GSSLFKRGYRTKGAIVKKNMAAIILOLSNWPDKPLIDPTGSGGTFCIEAVMIARK MAGLRRSFAFEENWISDRLIQEVRETAAKKVDRLELDIMGCDIDARMVEIAKANA QVAGVAGDITFKQMRVQDLRSKINGVIISNPYGERISDDAGYTKLVAEMGOVAPL KTWSKFILTSDEAFESKYGSQADKKRKLNGTLKVDLYQYFGQVRKROEVK"
gene	5514..6908
	/gene="Sp0374"
CDS	5514..6908
	/gene="Sp0374"
	/note="identified by Glimmer2; putative"
	/codon_start=1
	/transl_table=11
	/product="hypothetical protein"
	/protein_id="AAK74541.1"
	/db_xref="GI:14971843"
	/translation="MSKRRNRHKKGEQEPQDFDEAKELTVQAIRKNEEVESVLP EDSIDLVKQHRDEIADQATROIKKEEFVETQSDLDLOEMREAVEKSEASSEEV PSESDILPLDDEEGQLDPLLDENPTTEVEEQNLSRLDQEDSEKSKKSGF ILTVLAVSVIICVAYVYVQVARSKEIETQSQTANSDVDDEFNTLYDAFYDSDN KTKALNSQDKLSQLTKLLDKLEGSREHTLAKSKYDSLATQIKAIQDVNAOPEKPAIV DGVLTKAKASDAKFTDKTGTNDELVDKAIISLGKSOOTSTSSSSSOPTSSSSSS QASNTSEPKPSSSNETRSSEVNMGLSSAGAVORSASRAYNOSAIDDNSNAW PFADGVLEOILATSRSGYITGDQYILERVNIVNGYNYLKKPGTYTLFTLCKTGY FVNGAGHADDLDY"
gene	6984..8408
	/gene="Sp0375"
CDS	6984..8408
	/gene="Sp0375"
	/note="identified by match to TIGR protein family HMM TIGR00873"
	/codon_start=1
	/transl_table=11
	/product="6-phosphogluconate dehydrogenase, decarboxylating"
	/protein_id="AAK74542.1"
	/db_xref="GI:14971844"
	/translation="MTKANFGVYGMVAMGRNLALNIESRGYTVATYINRSKEDVIA CPEKNFVPSYDVSEFVNSIEKPRIMLMVQAGPGTDATIQALLPHLDKGDILDGGN TFYKDTIRRENLANSINGIFGTGYSGEGKALEGPGIMPQKQKAYELVADVLEIS AKAPEDGKPCVTYIGPDGAGHYVVMVHNGIEYDMOLIAESYDLMOHLGLLSAEDMAE IFTENWKGELSDYLIEITADILSRKDDQDGPVDYILDAAGNKGTKWTSQSSIDL GVPLSLITESVFARYISTYKEERVHASKVLPKPAAFNFEQDKAELEIKIRQALYFSKI ISYAQGFQALRVASKENNMNLPFADIASIWRDGCITIRSRFLOKITDAYNRDADLANLL LDEYFLDVTAKYQQAQRDIALAVQAGVPVPTFSAITFYDSYRSADLPANLIQAQRD YFGAHTYQKKKEGTFHYSWYDEK"
gene	8420..9109
	/gene="Sp0376"
CDS	8420..9109
	/gene="Sp0376"
	/codon_start=1
	/transl_table=11
	/product="DNA-binding response regulator"
	/protein_id="AAK74543.1"
	/db_xref="GI:14971845"
	/translation="MGRILLILEKERNLAHFLSLELOKEQYRVLDVEEGQKALSMAQ TDYDLMLNVLNGLDMMQDAEFKRSRTPASVIMILDHWDIQLQEELEVQVFAVYIY KPVLIENLVARISAFRCGRDFTDQHCILMKVPTRYNLRIDVHHQTVYRGEMIALTR REYDLLATMGSKKVLTRQELLLESVMKYESATETNIDVYIRYLSKLDLVKQKSYIK TVRGVGYTMQE"
gene	9208..10230
	/gene="Sp0377"
CDS	9208..10230
	/gene="Sp0377"
	/codon_start=1
	/transl_table=11

Db 1179 CGTTCACGATGACCCCTATAATTAACCTGGGACAAATACTCCTGTTTATAAATGGGATAG 1238
QY 1081 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAATCGCGAAAGCTCCACG 1140
Db 1239 GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACAATCGCGAAAGCTCCACG 1298
QY 1141 CGTGGAACTCTAAACAAGGTCGGACTCAACCGCGCAAGACTTTCCCTAAATGGTCTAGG 1200
Db 1299 CGTGGAACTCTAAACAAGGTCGGACTCAACCGCGCAAGACTTTCCCTAAATGGTCTAGG 1358
QY 1201 ATTCGACTACCAAGTATTCACCTACTCAATGCGCATTTCAAGTACACACCAACGATCAGA 1260
Db 1359 ATTCGACTACCAAGTATTCACCTACTCAATGCGCATTTCAAGTACACACCAACGATCAGA 1418
QY 1261 CAAAAAATATGAGCAAGTAGTGAAGAAGTGGCTGCTTACGCTTGGCTTTGGCAATGG 1320
Db 1419 CAAAAAATATGAGCAAGTAGTGAAGAAGTGGCTGCTTACGCTTGGCTTTGGCAATGG 1478
QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGAGTGAAAA 1380
Db 1479 TGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGATGGAGTGAAAA 1538
QY 1381 AGAGTCTCTAATGTCGGAACCTGCGCCATGAAGGAACGACAGCCTATATGATGACCGA 1440
Db 1539 AGAGTCTCTAATGTCGGAACCTGCGCCATGAAGGAACGACAGCCTATATGATGACCGA 1598
QY 1441 CATGATGAAACAGTCTTGACCTTATGGAACCTGGACGAATGCTATCTTGGCTGGCTCC 1500
Db 1599 CATGATGAAACAGTCTTGACCTTATGGAACCTGGACGAATGCTATCTTGGCTGGCTCC 1658
QY 1501 TCAGGCTGGTAAACAGGAACCTCTAATATACAGACGAGGAATTTGAAACACCATCAA 1560
Db 1659 TCAGGCTGGTAAACAGGAACCTCTAATATACAGACGAGGAATTTGAAACACCATCAA 1718
QY 1561 GACCTCTCAATTTGAGCACTGTGAACCTATTTGCTGCTATACGCGTAATATATCAAT 1620
Db 1719 GACCTCTCAATTTGAGCACTGTGAACCTATTTGCTGCTATACGCGTAATATATCAAT 1778
QY 1621 GCGTGTATGACAGGCTATTTACCGCTGTACACACCTTTGAGCAATGGCCTTACGGT 1680
Db 1779 GCGTGTATGACAGGCTATTTACCGCTGTACACACCTTTGAGCAATGGCCTTACGGT 1838
QY 1681 CCGTCCCAAGTTTACCGCTGTATGATGACCTATACCTGTCTGAAGGAAGCAATCCAGAGA 1740
Db 1839 CCGTCCCAAGTTTACCGCTGTATGATGACCTATACCTGTCTGAAGGAAGCAATCCAGAGA 1898
QY 1741 TTGGAATATACAGAGGCTCTACAGAAATGGAATTCGTATTTAAAAATGGTCTCG 1800
Db 1899 TTGGAATATACAGAGGCTCTACAGAAATGGAATTCGTATTTAAAAATGGTCTCG 1958
QY 1801 TTCTACGTGGAACCTCACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 1860
Db 1959 TTCTACGTGGAACCTCACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 2018
QY 1861 ATCAGATAGTTCACTTACAGCTTAGCTCAACCACTCCAGCAACAAATATATAGTACGAC 1920
Db 2019 ATCAGATAGTTCACTTACAGCTTAGCTCAACCACTCCAGCAACAAATATATAGTACGAC 2078
QY 1921 TACCAATCCTTACAAATATACCAACAATCAATACACCCCTGATCAACAAATCAGAA 1980
Db 2079 TACCAATCCTTACAAATATACCAACAATCAATACACCCCTGATCAACAAATCAGAA 2138
QY 1981 TCCTCAACAGCAGCAACCA 1999
Db 2139 TCCTCAACAGCAGCAACCA 2157

RESULT 5
LOCUS SPON8
DEFINITION S. pneumoniae (45607) ponA gene for penicillin-binding protein 1a.
ACCESSION X67873
VERSION X67873.1 GI:47419

penicillin-binding protein 1a; ponA gene.
Streptococcus pneumoniae.
Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2157)
Hakenbeck, R.
Direct Submission
Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.
Genetik, Ihnestr. 73, 1000 Berlin 33, FRG
2 (bases 1 to 2157)
Martin, C., Sibold, C. and Hakenbeck, R.
Relatedness of penicillin-binding protein 1a genes from different
clones of penicillin-resistant Streptococcus pneumoniae isolated in
South Africa and Spain
EMBO J. 11 (11), 3831-3836 (1992)
93010977 Location/Qualifiers
1..2157
/organism="Streptococcus pneumoniae"
/strain="45607"
/db_xref="taxon:1313"
/clone="pcMK1-pcMK16"
1..2157
/gene="ponA"
1..>2157
/gene="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="CAA48073.1"
/db_xref="GI:47420"
/translation="MNKPTILRLIKLSIFLSLVIAIVLGGVFFVYVSKAPSLSE
SKLVATSSKIYDNKNQIADLGSERRVNAQNDIPTLVKAIVSIEDHRRFPDHRGID
TIRLIGAFRLNLSQNSGSLTQQLIKLTFTSTSDQTSIRAKQAWLAIQLEQK
ATQGEILTYINKVNSNGNYGMQTAQNYGKDLNNLSLPALLAGMPAQNPQYDP
YSPHQAQDRRLNLSNKNQYISAEQYKAVNPTITDGLSKASNPAYMDNYL
KEVINQVEEETGYNLLTTGMDVYVNDQEAQKHLMDIYNTDEYVAYPDELYQVASTIV
DVSNCKYIAOLGARHOSNVSGFINOAVETNRDWSGTMKPTIDVAPALEYGVYDSTAT
IVHDEPNYPCNTNPVYNDRGYFNITLQVALQOSNRVPAVETLNKVLNRAKTFELN
GLGIDYPSIHYNSNAISSNTESDKKYGASSEKMAAAYAAFAANGTYIYKPMYIHKVYFS
DGSKEFNSVTRAMKETATYMTTMMKTVLTITGTRNAYLAWLPQAQKTGTSNTIDE
EIEHNIKTSQFVADELEFAGYTRYISMAVWTGYSNRLTPLYNGLTVAAKYVRSMTY
LSEGSNPEDWNIPEGLYRNGEFVFKNGARSTWSSPAQPPSTESSSSSSSSSTSQSS
STTPTSNNTTNNNTNQSNTPDQONQNPQAP"
BASE COUNT 696 a 520 c 411 g 530 t
ORIGIN

Query Match 79.6%; Score 1591; DB 1; Length 2157;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1991; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATAAAATCAACTCATTCGCTGCTGGTTCGAAACGCGGCTCAA 60
Db 159 TAAATCTACGACATAAAATCAACTCATTCGCTGCTGGTTCGAAACGCGGCTCAA 218
QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGCA 120
Db 219 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTCTATCGAAGCA 278
QY 121 TCGCTTCTTCGACACAGGGGATTTGATACCATCGTATCTCTGGAGCTTTCTTCGCGAA 180
Db 279 TCGCTTCTTCGACACAGGGGATTTGATACCATCGTATCTCTGGAGCTTTCTTCGCGAA 338
QY 181 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCAACCAACAGTTGATTAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCAACCAACAGTTGATTAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCGACGACGACTATTTCTCGTAGGCTCAGGAGCTTGGTT 300
Db 399 TTACTTTTCAACTTCGACTTCGACGACGACTATTTCTCGTAGGCTCAGGAGCTTGGTT 458

Qy	301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAATAA	360
Db	459	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAATAA	518
Qy	361	GGTCTACATGTCCTAATGGGAATATGGRAATCGAGACACGAGCTCAAAACTACTACTGTAA	420
Db	519	GGTCTACATGTCCTAATGGGAATATGGRAATCGAGACACGAGCTCAAAACTACTACTGTAA	578
Qy	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTCAGGCACC	480
Db	579	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTCAGGCACC	638
Qy	481	AAACCAATATGACCCCTATTACATCCAGAAGACGCCCAAGACGCCGGAACCTTGGTCTT	540
Db	639	AAACCAATATGACCCCTATTACATCCAGAAGACGCCCAAGACGCCGGAACCTTGGTCTT	698
Qy	541	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600
Db	699	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	758
Qy	601	ACCAATTACTGATGGACTCAAAGTCTCAAAATCAGCAAGTAATACCCCTGCTTACATGGA	660
Db	759	ACCAATTACTGATGGACTCAAAGTCTCAAAATCAGCAAGTAATACCCCTGCTTACATGGA	818
Qy	561	TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGAACAGGCTATAACCTACTCAC	720
Db	819	TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGAACAGGCTATAACCTGCTCAC	878
Qy	721	AACCTGGATGGATGCTACACAAATGTAGACCAAGACTCAAAAAATCTGTGGGATAT	780
Db	879	AACCTGGATGGATGCTACACAAATGTAGACCAAGACTCAAAAAATCTGTGGGATAT	938
Qy	781	TTACAATACAGAGAAATACGTTGCCTATCCAGACGATGAATTGCAAGTCGCTTACCAT	840
Db	939	TTACAATACAGAGAAATACGTTGCCTATCCAGACGATGAATTGCAAGTCGCTTACCAT	998
Qy	841	TGTTGATGTTCTTAACGGTAAAGTCAATGCCAGCTAGGACGCGCCATCAGTCAAGTAA	900
Db	999	TGTTGATGTTCTTAACGGTAAAGTCAATGCCAGCTAGGACGCGCCATCAGTCAAGTAA	1058
Qy	901	TGTTTCCTTCGGAATTAACCAAGCAGTAGAAACAACCGCAGCTGGGGATCAACTATGAA	960
Db	1059	TGTTTCCTTCGGAATTAACCAAGCAGTAGAAACAACCGCAGCTGGGGATCAACTATGAA	1118
Qy	961	ACCGATCACAGACTATGCTCTGCCTTGAGTACGGGTCTACGATTCACCTGCTACTAT	1020
Db	1119	ACCGATCACAGACTATGCTCTGCCTTGAGTACGGGTCTACGATTCACCTGCTACTAT	1178
Qy	1021	CGTTACAGATGAGCCCTATAACTACCTGGGACAATACTCCTGTTTATACTGGGATAG	1080
Db	1179	CGTTACAGATGAGCCCTATAACTACCTGGGACAATACTCCTGTTTATACTGGGATAG	1238
Qy	1081	GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGCGAACGTCGCCAGC	1140
Db	1239	GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACAATCGCGAACGTCGCCAGC	1298
Qy	1141	CGTGGAACTCTAAACAAGTCGGACTCAACCGGCCCAAGACTTTCCTAAATGGTCTAGG	1200
Db	1299	CGTGGAACTCTAAACAAGTCGGACTCAACCGGCCCAAGACTTTCCTAAATGGTCTCGG	1358
Qy	1201	AATCGACTACCCAAAGTATTCACTACTCAAAATGCCATTTCAAGTAAACAACCGAATCAGA	1260
Db	1359	AATCGACTACCCAAAGTATTCACTACTCAAAATGCCATTTCAAGTAAACAACCGAATCAGA	1418
Qy	1261	CAAAAAATATGGAGCAAGTAGTGAAGAATGSGTGTCTTACGCTGCCTTTGCCAAATGG	1320
Db	1419	CAAAAAATATGGAGCAAGTAGTGAAGAATGSGTGTCTTACGCTGCCTTTGCCAAATGG	1478
Qy	1321	TGGAAGTACTATAAACCAATGTATATCCATAAGTCGCTCTTTAGTCATGGGAGTGAAGA	1380
Db	1479	TGGAAGTACTATAAACCAATGTATATCCATAAGTCGCTCTTTAGTCATGGGAGTGAAGA	1538

Qy	1381	AGAGTTCCTAAATGTGTCGGAACTCGTGCCCATGAAGAAACGACAGCGCTATATGATGACCGA	1444
Db	1539	AGAGTTCCTAAATGTGCGAACTCGTGCCCATGAAGAAACGACAGCGCTATATGATGACCGA	1598
Qy	1441	CATGATCAAAACAGCTCTTGACCTATATGAACTTGACCAATGCTATCTTGCTTGGCTCCC	1500
Db	1599	AATGATCAAAACAGCTCTTGACCTATATGAACTTGACCAATGCTATCTTGCTTGGCTCCC	1658
Qy	1501	TCAGGCTGGTAAACACAGGAACCTCTTAACCTATACAGACGAGGAAATGAAACACCATCAA	1560
Db	1659	TCAGGCTGGTAAACACAGGAACCTCTTAACCTATACAGATGAGGAAATGAAACACCATCAA	1718
Qy	1561	GACCTCTCAATTTGTAGCACTGATGAACCTATTTGCTGGCTATACGCGTAATAATATCAAT	1620
Db	1719	GACCTCTCAATTTGTAGCACTGATGAACCTATTTGCTGGCTATACGCGTAATAATATCAAT	1778
Qy	1621	GGCTGTATGACAGGCTATCTTAACCGTCTGACACCACCTTGTAGGCAATGCGCTTACGGT	1680
Db	1779	GGCTGTATGACAGGCTATCTTAACCGTCTGACACCACCTTGTAGGCAATGCGCTTACGGT	1838
Qy	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA	1740
Db	1839	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTGAAGGAAGCAATCCAGAGA	1898
Qy	1741	TTGGAATATACCAAGAGGGGCTCTACAGAAATGGAGAATTCGTAATTTAAATAATGTTGCTCG	1800
Db	1899	TTGGAATATACCAAGAGGGGCTCTACAGAAATGGAGAATTCGTAATTTAAATAATGTTGCTCG	1958
Qy	1801	TTCTACGTGAACCTACCTGCTCCCAACAACCCCACTCAACTGGAAGCTTCAAGCTCATC	1860
Db	1959	TTCTACGTGAAGCTACCTGCTCCCAACAACCCCACTCAACTGGAAGCTTCAAGCTCATC	2018
Qy	1861	ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACCACTCCAAGCACAAATAATAGTAGCAC	1920
Db	2019	ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACCACTCCAAGCACAAATAATAGTAGCAC	2078
Qy	1921	TACCAATCCTTAACAATAATACGCAACAATAAATAACAAGCCCTGATCAACAAATACGAA	1980
Db	2079	TACCAATCCTTAACAATAATACGCAACAATAAATAACAAGCCCTGATCAACAAATACGAA	2138
Qy	1981	TCCTCAACGACGACACCA	1999
Db	2139	TCCTCAACGACGACACCA	2157
RESULT	6		
AF210746			
LOCUS	AF210746	2160 bp DNA linear	ECT 17-SEP-
DEFINITION	Streptococcus pneumoniae isolate SP 1513 penicillin-binding protein 1A (pbpia) gene, complete cds.		
ACCESSION	AF210746		
VERSION	AF210746.1		
KEYWORDS	GI:6563338		
SOURCE	Streptococcus pneumoniae.		
ORGANISM	Streptococcus pneumoniae.		
REFERENCE	1 (bases 1 to 2160)		
AUTHORS	Ferroni,A. and Berche,P.		
TITLE	Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children		
JOURNAL	J. Med. Microbiol. 50 (9), 828-832 (2001)		
MEDLINE	21432820		
PUBMED	11549185		
REFERENCE	2 (bases 1 to 2160)		
AUTHORS	Ferroni,A. and Berche,P.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-DEC-1999) Microbiology, Hospital		
FEATURES	Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France		
source	Location/Qualifiers		
	1. .2160		
	/organism="Streptococcus pneumoniae"		

		/serotype="23F"			
		/isolate="sp 1513"			
gene		/db_xref="taxon:1313"			
	1. .2160				
CDS		/gene="pbpla"			
	1. .2160				
	/gene="pbpla"				
	/codon_start=1				
	/transl_table=11				
	/product="penicillin-binding protein 1A"				
	/protein_id="AAFI7256.1"				
	/db_xref="GI:5563339"				
	/translation="MNKPTILRLIKYLSIFSLVIAAIVLGGGVFFVYVSKAPLSLSE SKLVATSSKIYDNKNOIADIPDLVRAIVSIEDHFRFDRHIGID TIRILGATLRNLSQSGSTLTQOLIKLVFSTSDQTSRKAQAWLAIOLEQK ATKORILTVYINKVYMSNGYKMOATQONYGKDLNLSLPOLASAGNPOAPOYDP YSHPEAODRNLVSEMNQGYISAEQYKAVNTPITDGLQSLKSSNYPAYMDNYL KEVINQVEETGNLLTTGMVDYINVDQEAQKHLMDIYNTDEIYVAPDELOQVASTIV DVSNGKVIQALGARHQSNNVSGFINQAVETNRDWSMKPITDVPALLEYGVYDSTAT IVHDEPNYPTNTFPVNMDRGYFNGNITLQYALQOSRNVPAVETLNKVLNRAKTFLLN GLGIDYPSIHYSNATSSNTTESDKYASSEKMAAAYAAFAANGTYIYKPMYIHKVVF DGSEKEFSNVGTRAMKETAYMMTEMKTVLSYGNRYLAWLPQACKTGTSNYVDE ELENHIKTSQVAPDELFAGYTRYKSYMAVWTFYSNRLPLVGNGLTVAAKYVRSMTY LSEGNPDWNIPEELRYNNGEFVFNKGARSTWNSPAPOPPFSTESSSSSSSSSTSSQSS STTSTNNSTNTNPNNTNQSNFTPDQONQPPQAP"				
BASE COUNT	700 a	518 c	410 g	532 t	
ORIGIN					
	Query Match	74.5%;	Score 1489;	DB 1;	Length 2160;
	Best Local Similarity	99.5%;	Pred. No. 0;		
	Matches 1989;	Conservative	0;	Mismatches	10; Indels 0; Gaps 0;
QY	1	TAATAATCTACGACAATAAAATCAACTCATTCGTGACTTGGTTCCTGAACGCGCGTCAA	60		
DB	159	TAATAATCTACGACAATAAAATCAACTCATTCGTGACTTGGTTCCTGAACGCGCGTCAA	218		
QY	61	TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	120		
DB	219	TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	278		
QY	121	TGCTTCTTCGACACAGGGGATGTATACATCGGTATCTTCGGAGCTTCTTCGCGAA	180		
DB	279	TGCTTCTTCGACACAGGGGATGTATACATCGGTATCTTCGGAGCTTCTTCGCGAA	338		
QY	181	TCTCAAGCAATTCCTCCAAAGGTGGATCAACTCTCACCCCAACAGTTGATTAACTTGAC	240		
DB	339	TCTCAAGCAATTCCTCCAAAGGTGGATCAACTCTCACCCCAACAGTTGATTAACTTGAC	398		
QY	241	TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	300		
DB	399	TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	458		
QY	301	AGCGATTTCAGTTAGACAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA	360		
DB	459	AGCGATTTCAGTTAGACAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA	518		
QY	361	GGTCTACATGCTAATGGGAATATGGAAATGCAGACAGAGCTCAAACTACTATGGTAA	420		
DB	519	GGTCTACATGCTAATGGGAATATGGAAATGCAGACAGAGCTCAAACTACTATGGTAA	578		
QY	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTAGCCTTGGTGGAAATGCCCTCAGGCACC	480		
DB	579	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTAGCCTTGGTGGAAATGCCCTCAGGCACC	638		
QY	481	AAACCAATATGACCCCTATTACATCCAGACAGCCCCAAGACCCGCGAAATCTGGTCTT	540		
DB	639	AAACCAATATGACCCCTATTACATCCAGACAGCCCCAAGACCCGCGAAATCTGGTCTT	698		
QY	541	ATCTGAAATGAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600		
DB	699	ATCTGAAATGAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	758		

QY	601	ACCAATTA	CTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTTGCCTTACATCGA	666	
DB	759	ACCAATTA	CTGATGGAGCTACAAAAGTCTCAAAATCAGCAAGTAATTAACCTTGCCTTACATCGA	818	
QY	661	TAATTA	ACCTCAAGSAGTCAATCAATCAAGTTGAAGAAGAACAGCGCTATAACCTTACTCAC	720	
DB	819	TAATTA	ACCTCAAGSAGTCAATCAATCAAGTTGAAGAAGAACAGCGCTATAACCTTACTCAC	878	
QY	721	AACTGGGAT	GGATGTCTTACAAAATGTAGACCAGAAAGCTCAAAAACATCTTGTGGGATAT	780	
DB	879	AACTGGGAT	GGATGTCTTACAAAATGTAGACCAGAAAGCTCAAAAACATCTTGTGGGATAT	938	
QY	781	TTACAATACAGACGAATACGTTGCCTATCCAGACGATGAATTGCAAGTTCGCTTCTACCAT	840		
DB	939	TTACAATACAGACGAATACGTTGCCTATCCAGACGATGAATTGCAAGTTCGCTTCTACCAT	998		
QY	841	TGTTGATGTTTCTAACCGGTAAAGTCATTGGCCAGCTAGGACGACGCCATCAGTCAAGTAA	900		
DB	999	TGTTGATGTTTCTAACCGGTAAAGTCATTGGCCAGCTAGGACGACGCCATCAGTCAAGTAA	1058		
QY	901	TGTTTCTCTCGGAATTAACCAAGCAGTAGAAAACAAACCGGACGTGGGATCAACATATGAA	960		
DB	1059	TGTTTCTCTCGGAATTAACCAAGCAGTAGAAAACAAACCGGACGTGGGATCAACATATGAA	1118		
QY	961	ACCGATCACAGACTATGCTCCTCGCTTGGAGTAGCGTGTCTACGATTCAACTGCTACTAT	1020		
DB	1119	ACCGATCACAGACTATGCTCCTCGCTTGGAGTAGCGTGTCTACGATTCAACTGCTACTAT	1178		
QY	1021	CGTTCCAGATGAGCCCTATAACTACCTGGGACAAAATAC	TCTCTGTTTATACTGGGATAG	1088	
DB	1179	CGTTCCAGATGAGCCCTATAACTACCTGGGACAAAATAC	CCCTGTTTATACTGGGATAG	1238	
QY	1081	GGCTACTTTGGCAACATCACTTGC	CAATACGCCTCGCAAAATCGGAAACGTCCCCAGC	1144	
DB	1239	GGCTACTTTGGCAACATCACTTGC	CAATACGCCTCGCAAAATCGGAAACGTCCCCAGC	1298	
QY	1141	CGTGGAACTCTAAACAAGTCGGACTCAACCGCGCCCAAGACTTTCC	TAAATGTC	1200	
DB	1299	CGTGGAACTCTAAACAAGTCGGACTCAACCGCGCCCAAGACTTTCC	TAAATGTC	1358	
QY	1201	AATCGACTACCCAAAGTATTCAC	TACTCAAAATGCCATTTCAAGTAACACACCGAATCAGA	1266	
DB	1359	AATCGACTACCCAAAGTATTCAC	TACTCAAAATGCCATTTCAAGTAACACACCGAATCAGA	1418	
QY	1261	CAAAAATATGGAGCAAGTAGTGA	AAAAAGATGGCTGCTTACCGCTGCTTTGCAAAATGG	1322	
DB	1419	CAAAAATATGGAGCAAGTAGTGA	AAAAAGATGGCTGCTTACCGCTGCTTTGCAAAATGG	1478	
QY	1321	TGGAACCTTACTATAAACCAATGTATATCCATAAAGTCGTCTTTAGT	GATGGAGTGA	1386	
DB	1479	TGGAACCTTACTATAAACCAATGTATATCCATAAAGTCGTCTTTAGT	GATGGAGTGA	1538	
QY	1381	AGAGTTCTCTAATGTCGGAAC	TGTCGCATGAAGGAAACGACGCCCTATATGATGACCGA	1440	
DB	1539	AGAGTTCTCTAATGTCGGAAC	TGTCGCATGAAGGAAACGACGCCCTATATGATGACCGA	1598	
QY	1441	CATGATGAAACAGTCTTGAC	TTATGGAAC	TGGACGAAATTCCTTGCCTTGGCTCCC	1500
DB	1599	AATGATGAAACAGTCTTGAC	TTATGGAAC	TGGACGAAATTCCTTGCCTTGGCTCCC	1658
QY	1501	TCAGGCTGGTAAACAGGAACTCT	TAACATACAGACGAGGAAATTTGAAACACCATCAA	1566	
DB	1659	TCAGGCTGGTAAACAGGAACTCT	TAACATACAGATGAGGAAATTTGAAACACCATCAA	1718	
QY	1561	GACCTCTCAATTTGTAGC	ACCTGATTAATTTGCTGGCTATACGCGTAAATATTC	1622	
DB	1719	GACCTCTCAATTTGTAGC	ACCTGATTAATTTGCTGGCTATACGCGTAAATATTC	1778	
QY	1621	GGCTGTATGACAGGCTATTTCT	AACCGTCTGACACCACTTGTAGGCAATG	1686	
DB	1779	GGCTGTATGACAGGCTATTTCT	AACCGTCTGACACCACTTGTAGGCAATG	1838	
QY	1681	CGCTGCAAAAGTTTACCGCTCT	ATGATGACCTACCTCTGCTGAAGGAAGCAATCCAGAA	1744	

QY	601	ACCAATTAATGATGGACTACAAAGTCTCAAAATAGCAAGTAATATACCTGCTTACATGA	660
DB	759	ACCAATTAATGATGGACTACAAAGTCTCAAAATAGCAAGTAATATACCTGCTTACATGA	818
QY	661	TAATTAACCTCAAGGAAGTCAATCAAGTTGAAGGAAGCAAGGCTATAACCTACTCAC	720
DB	819	TAATTAACCTCAAGGAAGTCAATCAAGTTGAAGGAAGCAAGGCTATAACCTACTCAC	878
QY	721	AACTGGGATGGATGCTCACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
DB	879	AACTGGGATGGATGCTCACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	938
QY	781	TTACAATACAGACGAATACGTTGCCCTATCCAGACGATGAATGTCAAGTCCGTTCTACCAT	840
DB	939	TTACAATACAGACGAATACGTTGCCCTATCCAGACGATGAATGTCAAGTCCGTTCTACCAT	998
QY	841	TGTTGATGTTTCTAAGCGTTAAAGTCAATGCCAGCTAGGACGACGCCATCAGTCAAGTAA	900
DB	999	TGTTGATGTTTCTAAGCGTTAAAGTCAATGCCAGCTAGGACGACGCCATCAGTCAAGTAA	1058
QY	901	TGTTTCTTCCGGAATTAACCAAGCAGTACAAACAAACCGGACTGGGGATCAACTATGAA	960
DB	1059	TGTTTCTTCCGGAATTAACCAAGCAGTACAAACAAACCGGACTGGGGATCAACTATGAA	1118
QY	961	ACCGATACAGACTATGCTCTTCCCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT	1020
DB	1119	ACCGATACAGACTATGCTCTTCCCTTGGAGTACGGTGTCTACGATTCAACTGCTACTAT	1178
QY	1021	CGTTCACGATGAGCCCTATAACTACCTGGGCAAAATACCTCTCTGTTTATAACTGGGATAG	1080
DB	1179	CGTTCACGATGAGCCCTATAACTACCTGGGCAAAATACCTCTCTGTTTATAACTGGGATAG	1238
QY	1081	GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACCAATCGGAAACGCTCCAGC	1140
DB	1239	GGGCTACTTTGGCAACATCACCTTGAATACGCCCTGCAACCAATCGGAAACGCTCCAGC	1298
QY	1141	CGTGGAAACTCTAAACAAGTGGAGCTCAACCGCGCAAGACTTTCTTAAATGGTCTAGG	1200
DB	1299	CGTGGAAACTCTAAACAAGTGGAGCTCAACCGCGCAAGACTTTCTTAAATGGTCTAGG	1358
QY	1201	AATCGACTACCAAGTATTTCACCTACTCAATGCCATTTCAAGTAAACACACCAACATCAGA	1260
DB	1359	AATCGACTACCAAGTATTTCACCTACTCAATGCCATTTCAAGTAAACACCAACATCAGA	1418
QY	1261	CAAAAAATATGAGCAAGTAGTGAAAAGATGGCTGCTGTCTAGCTGCTCTTTGCAAAATGG	1320
DB	1419	CAAAAAATATGAGCAAGTAGTGAAAAGATGGCTGCTGTCTAGCTGCTCTTTGCAAAATGG	1478
QY	1321	TGGAACCTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTGTAGTGGGAGTGAATA	1380
DB	1479	TGGAACCTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTGTAGTGGGAGTGAATA	1538
QY	1381	AGAGTTCTTATATGTCGGAACCTCGTCCCATGAAGGAACGACAGCCCTATATGATGACCA	1440
DB	1539	AGAGTTCTTATATGTCGGAACCTCGTCCCATGAAGGAACGACAGCCCTATATGATGACCA	1598
QY	1441	CATGATGAAACAGCTTCTGACTTATGGAATGAGCAAAATGCTATCTTGTCTGGCTCCC	1500
DB	1599	CATGATGAAACAGCTTCTGACTTATGGAATGAGCAAAATGCTATCTTGTCTGGCTCCC	1558
QY	1501	TCAGGCTGTGTAACAGGAACCTCTAACTATACAGACGAGGAATTTGAAACCAACATCAA	1560
DB	1659	TCAGGCTGTGTAACAGGAACCTCTAACTATACAGATGAGGAATTTGAAACCAACATCAA	1718
QY	1561	GACCTCTCAATTTGTAGCAGCTGTAGCACTNTTGTGCTGCTATACCGGTAAATATTCAT	1620
DB	1719	GACCTCTCAATTTGTAGCAGCTGTAGCACTNTTGTGCTGCTATACCGGTAAATATTCAT	1778
QY	1621	GCTGTATGACAGGCTATTCTAAACGCTGTGACACCACTTTGTAGGCAATGGGCTTACGGT	1680
DB	1779	GCTGTATGACAGGCTATTCTAAACGCTGTGACACCACTTTGTAGGCAATGGGCTTACGGT	1838
QY	1681	CGCTGCCAAAGTTTACCCTCTATGATGACCTACCTGTCTGTAAGGAAGCAATCCAGAA	1740

Db 1839 CCGTCGCAAGTTACCGCTGATGACCTATCTGTCTGAAGGAAGCAATCCAGAGGA 1898
Qy 1741 TTGGAATATACAGAGGGCTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1800
Db 1899 TTGGAATATACAGAGGACTCTACAGAAATGGAGAAATTCGTATTTAAAAATGGTCTCG 1958
Qy 1801 TTCTAGTGGAACTACCTGCTCCACACAAACCCCATCAACTGAAAGTTCAAGCTCATC 1860
Db 1959 TTCTAGTGGAACTACCTGCTCCACACAAACCCCATCAACTGAAAGTTCAAGCTCATC 2018
Qy 1861 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAAGCACAATAAATAGTAGCAG 1920
Db 2019 ATCAGATAGTTCAACTTCACAGTCTAGCTCAACCACTCAAGCACAATAAATAGTAGCAG 2078
Qy 1921 TACCAATCCTAACAATAATACGCAACAATCAATAACAACCCCTGATCAACAAAATCAGAA 1980
Db 2079 TACCAATCCTAACAATAATACGCAACAATCAATAACAACCCCTGATCAACARAATCAGAA 2138
Qy 1981 TCCTCAACCAAGCACAACCA 1999
Db 2139 TCCTCAACCAAGCACAACCA 2157

RESULT 7
SPON7
LOCUS SPON7 2157 bp DNA linear BCT 11-MAY-1995
DEFINITION S.pneumoniae (63915) ponA gene for penicillin-binding protein 1a.
ACCESSION X67872
VERSION X67872.1 GI:47417
KEYWORDS penicillin-binding protein 1a; ponA gene.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 2157)
AUTHORS Hakenbeck, R.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.
Genetik, Ihnestr. 73, 1000 Berlin 33, FRG
REFERENCE 2 (bases 1 to 2157)
AUTHORS Martin, C., Sibold, C. and Hakenbeck, R.
TITLE Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain
JOURNAL EMBO J. 11 (11), 3831-3836 (1992)
MEDLINE 93010977
FEATURES
source
1..2157 Location/Qualifiers
/organism="Streptococcus pneumoniae"
/strain="63915"
/db_xref="taxon:1313"
/clone="pcMK1-pcMK16"
1..2157
/gene="ponA"
1..>2157
/gene="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="CAA48072.1"
/db_xref="GI:47418"
/db_xref="SWISS-PROT:Q04707"
translation="MNKPTILRLIKYLSFLSLVIAIVLGGVFFVYVSKAPLSLSE
SKLVATSSKIVDNKLIADLSERRVNAOANDIPTDLVRAIVSIEDHRFPDRHGD
TIRLGAFLRNQSLGGSTLQOLIKLAYESTSDQIRSKRAQEAWLAIQLEQK
ATKQELIYYINKVYNSNGTGMQTAQNYVKGDLNLSLPOLALLAGMPQNPYIDP
YSHPPAARDNRNLVSEMKNOGYISAEQYERKAVNPITDGLQSLKSASNPAYMDNYL
KEVINQVEETGYNLLTTGMDVYTNVDOEAQKHLMDIYNTDEYVAYPDELEQVASTIV
DVSNGKVALQGLARHQSNSVFGINQAVETNRDWSMTKPTIDYAPALEYGYIDSTAV
IVHDKPNYPGNTPVYNDRGYFGNITLOYALQOSNRVPAVELINKVLPNKRFLN
GLGDYPSIHVSNALSSNTTSDKKYGSSEKMAAAYAFANGGTYIYKPMYIHKVYFS
DGSEKFSNVGTRAIKETATYMTMDMKTVLSYGTGRNAYLAWLPQAGKTGTSNTIDE

ETENHIKTSQFVAPDELFAGYTRKYSMVWVTGYSNRLPLVGNGLUTVAAKYVRSMYTY
LSEGNPNEDWNIPEGLYRNGFEVFKNGSTWNSPAQPPSPSTESSSSDSTSQSS
STTPSTNNSTTTNNNTTQOSNTTDPQONONPPAPQ*
BASE COUNT 699 a 521 c 407 g 530 t
ORIGIN
Query Match 71.9%; Score 1438; DB 1; Length 2157;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 1 TAAATCTAGGACAATAAAAAATCAACTCATTTGCTGACTTGGGTCTCAAGCCGCGTCAA 60
Db 159 TAAATCTAGGACAATAAAAAATCAACTCATTTGCTGACTTGGGTCTCAAGCCGCGTCAA 218
Qy 61 TGCCCAAGCTAATGATATTCCCAACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGCCCAAGCTAATGATATTCCCAACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
Qy 121 TCGCTTCTTCGACCACAGGGGGATTGATCCATCCGTATCTCTGGAGCTTTCTTGGCGAA 180
Db 279 TCGCTTCTTCGACCACAGGGGGATTGATCCATCCGTATCTCTGGAGCTTTCTTGGCGAA 338
Qy 181 TCTGCAAGCAATTCCTCCAAAGGTGGATCAACTCTCACCACACAGTTGATTAAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCAAAGGTGGATCAACTCTCACCACAGTTGATTAAAGTTGAC 398
Qy 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 300
Db 399 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 458
Qy 301 AGGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGCCTACTATATAATAA 360
Db 459 AGGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGCCTACTATATAATAA 518
Qy 361 GGTCTACATCTCTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATGTTAA 420
Db 519 GGTCTACATCTCTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATGTTAA 578
Qy 421 AGACCTCAATAATTTAAGTTTACCTCAGTAGCTTGGCTGGGATGGAATGCTCAGCACACC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTAGCTTGGCTGGGATGGAATGCTCAGCACACC 638
Qy 481 AACCATAATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCCGGAAACTTGTCTT 540
Db 639 AACCATAATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCCGGAAACTTGTCTT 698
Qy 541 ATCTGAATCAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC 600
Db 699 ATCTGAATCAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC 758
Qy 601 ACCAATTACTGATGGACTACAAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGA 660
Db 759 ACCAATTACTGATGGACTACAAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGA 818
Qy 661 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAGAAACAGGCTATAACCTACTCAC 720
Db 819 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAGAAACAGGCTATAACCTACTCAC 878
Qy 721 AACTGGGATGGATGCTACACAAATCTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
Db 879 AACTGGGATGGATGCTACACAAATCTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 938
Qy 781 TTACAATACAGCAATACGTTCCCTATCCAGACGATGAATTCAGAGTCGCTTCTACCAT 840
Db 939 TTACAATACAGCAATACGTTCCCTATCCAGACGATGAATTCAGAGTCGCTTCTACCAT 998
Qy 841 TGTGTATGTTCTTAACGGTAAAGTCAATGCCCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
Db 999 TGTGTATGTTCTTAACGGTAAAGTCAATGCCCCAGCTAGGAGCAGCCATCAGTCAAGTAA 1058
Qy 901 TGTTCCTTCGGAAATTAACCAAGCAGTAGAACAACCCGAGCTGGGGATCAACTATGA 960

Db 1059 TGTTCCTCTCGGAATTAACCAAGCAGTAGAAGAAACAAACCGGACTGGGGATCAACTATGAA 1118
QY 961 ACCGATCACAGACTATGCTTCCTGGCTTGGAGTACGGTGTCTAGGATTCAACTGCTACTAT 1020
Db 1119 ACGGATCACAGACTATGCTTCCTGGCTTGGAGTACGGTGTCTAGGATTCAACTGCTACTAT 1178
QY 1021 CGTTCACGATGAGCCCTATTAACCTACCTGGGCAATACTCTGTTTATTAACCTGGGATAG 1080
Db 1179 CGTTCACGATGAGCCCTATTAACCTACCTGGGCAATACTCTGTTTATTAACCTGGGATAG 1238
QY 1081 GGGCTACTTTGGCAACATCACCTTGAATACGGCCCTGGCAACATCCGGAAGCTGCCAGC 1140
Db 1239 GGGCTACTTTGGCAACATCACCTTGAATACGGCCCTGGCAACATCCGGAAGCTGCCAGC 1298
QY 1141 CGTGGAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1200
Db 1299 CGTGGAACTCTAAACAAGTGGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAGG 1358
QY 1201 AATCGACTACCAAGTATTTCACTACTCAATCAATGCCATTTTCAAGTAAACACACCGGAATCAGA 1260
Db 1359 AATCGACTACCAAGTATTTCACTACTCAATCAATGCCATTTTCAAGTAAACACACCGGAATCAGA 1418
QY 1261 CAATAAATATGAGCAAGTAGTGAAGAATGGCTGCTGCTTACGCTGCCTTTGCAAAATGG 1320
Db 1419 CAATAAATATGAGCAAGTAGTGAAGAATGGCTGCTGCTTACGCTGCCTTTGCAAAATGG 1478
QY 1321 TGGAACTTACTATAAACCAATGTATATCCATAAGTCGCTTTAGTGATGGGAGTCAAAA 1380
Db 1479 TGGAACTTACTATAAACCAATGTATATCCATAAGTCGCTTTAGTGATGGGAGTCAAAA 1538
QY 1381 AGAGTTCTCTAATGTGCGGAACCTGTCGCATGAAGGAACGACAGCCATATATGATGACCGA 1440
Db 1539 AGAGTTCTCTAATGTGCGGAACCTGTCGCATGAAGGAACGACAGCCATATATGATGACCGA 1598
QY 1441 CATGATGAACAGCTTTCGACTTATGGAATGGAGTGGACCAATGCTATCTTGGCTGGCTCC 1500
Db 1599 CATGATGAACAGCTTTCGACTTATGGAATGGAGTGGAGTGGCTATCTTGGCTGGCTCC 1658
QY 1501 TCAGGCTGTGTAACAGGAACCTCTCACTATACAGACGAGGAATTTGAAACACACATCAA 1560
Db 1659 TCAGGCTGTGTAACAGGAACCTCTCACTATACAGACGAGGAATTTGAAACACACATCAA 1718
QY 1561 GACCTCTCAATTTGTAGCACTGTATGAATTTTGTGCTGCTATACCGTAAATATTTCAAT 1620
Db 1719 GACCTCTCAATTTGTAGCACTGTATGAATTTTGTGCTGCTATACCGTAAATATTTCAAT 1778
QY 1621 GGTGTATGACAGGCTATTTCAACGCTGTACACCACTTTGTAGGCAATGGCTTACGGT 1680
Db 1779 GGTGTATGACAGGCTATTTCAACGCTGTACACCACTTTGTAGGCAATGGCTTACGGT 1838
QY 1681 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACCTGCTGTAAGGAAGCAATCCAGAAGA 1740
Db 1839 CGCTGCAAAAGTTTACCGCTCTATGATGACCTATCTGTCTGAAGGAAGCAATCCAGAAGA 1898
QY 1741 TTGGAATATACGAGGGGCTCTACAGAAATGGAGAAATTCGTAATTTAAATATGGTCTCG 1800
Db 1899 TTGGAATATACGAGGGGCTCTACAGAAATGGAGAAATTCGTAATTTAAATATGGTCTCG 1958
QY 1801 TTCTAGTGAACCTACCTGCTCCACAACACCCCATCACTGAAAGTTCAAGCTTCATC 1860
Db 1959 TTCTAGTGAACCTACCTGCTCCACAACACCCCATCACTGAAAGTTCAAGCTTCATC 2018
QY 1861 ATCAGATAGTTCAACTTACAGCTAGCTACACCTCCAAAGCACAATAATAGTACGAC 1920
Db 2019 ATCAGATAGTTCAACTTACAGCTAGCTACACCTCCAAAGCACAATAATAGTACGAC 2078
QY 1921 TACCATCTACAAATATACGGAACAATCAATAACACCCCTGATCAACAAATCAGAA 1980
Db 2079 TACCATCTACAAATATACGGAACAATCAATAACACCCCTGATCAACAAATCAGAA 2138
QY 1981 TCCTCAACCCAGCACACCA 1999
Db 2139 TCCTCAACCCAGCACACCA 2157

RESULT 8

AX111315
LOCUS AX111315 2160 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2048 from Patent WO0123604.
ACCESSION AX111315
VERSION AX111315.1 GI:13927607
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Bergeron, M.G., Boissinot, M., Huletsky, A., m Nard, C., Ouellette, M.,
Picard, F.J. and Roy, P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2048 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
Location/Qualifiers
source
1..2160
/organism="Streptococcus pneumoniae"
/strain="R6"
/db_xref="taxon:1313"
BASE COUNT 697 a 522 c 413 g 528 t
ORIGIN

Query Match

Best Local Similarity 71.9%; Score 1438; DB 6; Length 2160;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATAAATAAATACTGCTGACTTGGGTTCTGAAGCGCGCTCAA 60
Db 159 TAAATCTAGCAATAAATAAATACTGCTGACTTGGGTTCTGAAGCGCGCTCAA 218
QY 61 TGCCCAAGCTAATGATATTTCCCAAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGCCCAAGCTAATGATATTTCCCAAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TGCCTTCTCGACACAGGGGATTTGATACCATCCGTATCTCTGAGGAGCTTCTTTCGCGCA 180
Db 279 TGCCTTCTCGACACAGGGGATTTGATACCATCCGTATCTCTGAGGAGCTTCTTTCGCGCA 338
QY 181 TCTGCAAGCAATTCCTCCAAAGTGGATCACTCTACCCACACAGTTGATTAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCAAAGTGGATCACTCTACCCACACAGTTGATTAAGTTGAC 398
QY 241 TTACTTTTCACTTCGACTTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 TTACTTTTCACTTCGACTTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTGATTAGACAAAAGCAACCAAGCAAGCAAAATCTTGACCTACTATATAATAA 360
Db 459 AGCGATTGATTAGACAAAAGCAACCAAGCAAGCAAAATCTTGACCTACTATATAATAA 518
QY 361 GGTCTACATGTCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAAACTACTATGGTAA 420
Db 519 GGTCTACATGTCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAAACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGTGGTGAATGCCCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGTGGTGAATGCCCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGACAGCCCAAGACCCCGGAACTTGTCTT 540
Db 639 AAACCAATATGACCCCTATTCACATCCAGACAGCCCAAGACCCCGGAACTTGTCTT 698
QY 541 ATCTGAAATGAAAAATCAAGGCTTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAAATGAAAAATCAAGGCTTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758

QY 601 ACCAATTACTGAGTACTCAAAAGTCTCAAAATCAGCAAGTAATATACCTGCTTACATGGA 660
|||||
Db 759 ACCAATTACTGATGGGCTACAAAGTCTCAAAATCAGCAAGTAATATACCTGCTTACATGGA 818
|||||
QY 661 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAGAAACAGGCTATACCTACTCAC 720
|||||
Db 819 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAGAAACAGGCTATACCTACTCAC 878
|||||
QY 721 AACTGGGATGGATGTTTACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
|||||
Db 879 AACTGGGATGGATGTTTACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 938
|||||
QY 781 TTCAATACAGACGAATACCTGGCTATCCAGACGATGAATGCAAGTGCCTTCTTACCAT 840
|||||
Db 939 TTCAATACAGACGAATACCTGGCTATCCAGACGATGAATGCAAGTGCCTTCTTACCAT 998
|||||
QY 841 TGTGTGATGTTTCAACGGTAAAGTCATTTGCCAGCTAGGAGCAGCCCATCAGTCAAGTAA 900
|||||
Db 999 TGTGTGATGTTTCAACGGTAAAGTCATTTGCCAGCTAGGAGCAGCCCATCAGTCAAGTAA 1058
|||||
QY 901 TGTGTTCTTCGGAATTAAACCAAGCAGTAGAAACAAACCCGAGCTGGGGATCAACTATGAA 960
|||||
Db 1059 TGTGTTCTTCGGAATTAAACCAAGCAGTAGAAACAAACCCGAGCTGGGGATCAACTATGAA 1118
|||||
QY 961 ACCGATCACAGACTATGCTCCTCGCTTGGAGTAGGAGTGTCTACGATTCACCTGCTACTAT 1020
|||||
Db 1119 ACCGATCACAGACTATGCTCCTCGCTTGGAGTAGGAGTGTCTACGAGTCAACTGCGCACTAT 1178
|||||
QY 1021 CGTTACGATGAGCCCTATAACTACCTCGGACAAATACTCCTGTTTATAACTGGGATAG 1080
|||||
Db 1179 CGTTACGATGAGCCCTATAACTACCTCGGACAAATACCCTGTTTATAACTGGGATAG 1238
|||||
QY 1081 GGCGTACTTTGGCAACATCACTTGGCAATACGCCCTGCACAAATCGCGAATCGTCCAGC 1140
|||||
Db 1239 GGCGTACTTTGGCAACATCACTTGGCAATACGCCCTGCACAAATCGCGAATCGTCCAGC 1298
|||||
QY 1141 CGTGGAACTCTAAACAGTCGGACTCAACCGGCCCAAGACTTTCCTAAATGGTCTAGG 1200
|||||
Db 1299 CGTGGAACTCTAAACAGTCGGACTCAACCGGCCCAAGACTTTCCTAAATGGTCTCGG 1358
|||||
QY 1201 AATCGACTACCCAAAGTATTCACACTACCAATGCCATTCAGTAACCAACCGCAATCAGA 1260
|||||
Db 1359 AATCGACTACCCAAAGTATTCACACTACCAATGCCATTCAGTAACCAACCGCAATCAGA 1418
|||||
QY 1261 CAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGCTGTACGCTGCCTTTGCAAAATGG 1320
|||||
Db 1419 CAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGCTGTACGCTGCCTTTGCAAAATGG 1478
|||||
QY 1321 TGGAACTTACTATAACCAATGATATCCATAAAGTGGCTTTAGTGATGGGAGTGAATA 1380
|||||
Db 1479 TGGAACTTACTATAACCAATGATATCCATAAAGTGGCTTTAGTGATGGGAGTGAATA 1538
|||||
QY 1381 AGAGTTCTCTAAATGTCGGAATCGTGCAATGAAGGAACACAGCCCTATATGATGACCGA 1440
|||||
Db 1539 AGAGTTCTCTAAATGTCGGAATCGTGCAATGAAGGAACACAGCCCTATATGATGACCGA 1598
|||||
QY 1441 CATGATGAAACAGTCTTGACTATGGAAGTGGACGAATGCCTATCTTGCTTGGCTCCC 1500
|||||
Db 1599 CATGATGAAACAGTCTTGACTATGGAAGTGGACGAATGCCTATCTTGCTTGGCTCCC 1658
|||||
QY 1501 TCAGGCTGGTAAACAGGAACCTCTTAATATACAGACGAGGAATGAAACACACATCAA 1560
|||||
Db 1659 TCAGGCTGGTAAACAGGAACCTCTTAATATACAGACGAGGAATGAAACACACATCAA 1718
|||||
QY 1561 GACCTCTCAATTTGTAGCACCTGATGACTATTTGCTGGCTATACGGTAAATATTCAT 1620
|||||
Db 1719 GACCTCTCAATTTGTAGCACCTGATGACTATTTGCTGGCTATACGGTAAATATTCAT 1778
|||||
QY 1621 GCGTGTATGACAGGCTATTTCAACCGCTGACACCACTGTGAGGCAATGSCCTTACGGT 1680
|||||
Db 1779 GCGTGTATGACAGGCTATTTCAACCGCTGACACCACTGTGAGGCAATGSCCTTACGGT 1838
|||||
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTGTGAAGGAAGCAATCCAGAAGA 1740
|||||

Db 1839 CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGGAAGCAATCCAGAGA 1898
|||||
QY 1741 TTGGAATATACCAGAGGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGCTGCTCG 1800
|||||
Db 1899 TTGGAATATACCAGAGGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGCTGCTCG 1958
|||||
QY 1801 TTCTACGTGGAACCTCACCTGCTCCACAAACACCCCATCAACTGAAAGTTTCAAGCTCATC 1860
|||||
Db 1959 TTCTACGTGGAACCTCACCTGCTCCACAAACACCCCATCAACTGAAAGTTTCAAGCTCATC 2018
|||||
QY 1861 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACCACTCCCAAGCACAAATATAGTAGCAC 1920
|||||
Db 2019 ATCAGATAGTTCAACTTTCACAGTCTAGCTCAACCACTCCCAAGCACAAATATAGTAGCAC 2078
|||||
QY 1921 TACCAATCCTTAACAATAATACGCAACAATCAATACACCCCTCATCAACAAAAATCAGAA 1980
|||||
Db 2079 TACCAATCCTTAACAATAATACGCAACAATCAATACACCCCTCATCAACAAAAATCAGAA 2138
|||||
QY 1981 TCCTCAACACGACACAACCA 1999
|||||
Db 2139 TCCTCAACACGACACAACCA 2157
|||||
RESULT 9
STRPONA
LOCUS
DEFINITION STRPONA 3378 bp DNA linear BCT 26-APR-1993
Streptococcus pneumoniae penicillin-binding protein (ponA) gene,
complete cds.
ACCESSION M90527.1 GI:153766
VERSION M90527.1 GI:153766
KEYWORDS penicillin-binding protein; ponA gene.
SOURCE Streptococcus pneumoniae (strain R6) DNA.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 3378)
AUTHORS Martin,C., Briese,T. and Hakenbeck,R.
TITLE Nucleotide sequences of genes encoding penicillin-binding proteins
from Streptococcus pneumoniae and Streptococcus oralis with high
homology to Escherichia coli penicillin-binding proteins 1A and 1B
J. Bacteriol. 174, 4517-4523 (1992)
JOURNAL
MEDLINE
FEATURES
source Location/Qualifiers
1. .3378
/organism="Streptococcus pneumoniae"
/strain="R6"
/db_xref="taxon:1313"
315. .3105
/gene="ponA"
315. .320
/gene="ponA"
353. .949
/gene="ponA"
/function="unknown"
/note="ORF"
/codon_start=1
/transl_table=11
/protein_id="AA26955.1"
/db_xref="GI:153767"
/translation="MVNYPHKVSSQRKQTSLSQPKNFANRGMSEFKINATINDYILSQ
GLAIVHKPTPIQIVQVDYPPORSRAKIVFRAQSTIDYSGVNGYVYIDFEVKETQK
KRAIPKMFPHQIOHMEQVLAQGGICFVLLHPFSQOETVLLPAFDLIRFYHQDKGQK
SMPLVIREYGVYEIKAGAFPQIPVNLVIREHLLGGKTR"
946. .3105
/gene="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein"
/protein_id="AA26956.1"
/db_xref="GI:153768"
/translation="MNKPTILRLIKYLSIFSLVIAAIVLGGGVFFYVSKAPSLSE
SKLVATSSSKIYDNKNQLIADLGSRRVNAQANDIPTDLVKAIVSIEDHRFDFHRGID

TIRILGAFUFLRNLOSLSQGSALTQOLIKLTVFSTSTSDTISRKAQEAWLAIQLEQK
ATKQELTIIYINKVYMSNGNYGKDLNLSLPOLALLAGMPQAYNDYP
YGRPAADORRLNVLSEMKNOGYISAEQYKAVNPTIDGLOSLKASNYPAIMONYL
KEVINQVEETGVNLLTTCMDVYTNVDOEAQKHLNDIYNTEDEVVAPDDELOVASTIV
DYSNGVIAQLGARQSSNVSEFINQAVETNRDWSGTHKPIITDYAPALEYGYVESTAT
IVHDEPYNTPGNTPIYNNDRGYFNITLOIALQOSRNPVAVETLNKVLGNRAKTFLN
GLGIDPYSNYSNAISSNTTESDKKYGASSEKMAAAYAFANGGYIYKPMYTHKVVFS
DGSERFNVSGTRAMKETAYNMKTDMKTVLSYGTGRNAYLAWLPQAGKTGTSNYTDE
ETENHIKTSOFVAPDELPAFYTRKYSMVAWMTGYSNRLTPLVNGLITVAAKVYRSMWTY
LSEGSNPEDMNIPEGLYRNGFEVFNKNGASTWSPPAQPPSTESSSSSDSSTSOSS
STTPSINNSTTNNPNNNTQOOSNTTDDOQNQNPQAP"

BASE COUNT 1093 a 783 c 609 g 893 t
ORIGIN

Query Match		71.9%;	Score 1438;	DB 1;	Length 3378;
Best Local Similarity		99.4%;	Pred. No. 0;		
Matches 1988;		Conservative 0;	Mismatches 11;	Indels 0;	Gaps 0;
QY	1	TAAATCTAGCACAATAAATAATCAATTCATTGCTGACTTGGGTCTTGAACGCCGCGTCAA	60		
DB	1104	TAAATCTAGCACAATAAATAATCAATTCATTGCTGACTTGGGTCTTGAACGCCGCGTCAA	1163		
QY	61	TGCCCAAGCTAATGATATCCCAACAGATTGGTTAAGGCAATCGTTCTATCCGAAGACCA	120		
DB	1164	TGCCCAAGCTAATGATATCCCAACAGATTGGTTAAGGCAATCGTTCTATCCGAAGACCA	1223		
QY	121	TCGCTTCTTCGACACACAGGGGGATTGATACCATCCGTATCCTGGAGCTTCTTTCGGCAA	180		
DB	1224	TCGCTTCTTCGACACACAGGGGGATTGATACCATCCGTATCCTGGAGCTTCTTTCGGCAA	1283		
QY	181	TCGCAAGCAATTCCTCCCAAGGTGGATCACTCAACCACACAGTTGATTAAAGTTGAC	240		
DB	1284	TCGCAAGCAATTCCTCCCAAGGTGGATCAGCTCACTCAACAGTTGATTAAAGTTGAC	1343		
QY	241	TTACTTTTCAACTTCGACTTCCACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	300		
DB	1344	TTACTTTTCAACTTCGACTTCCACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	1403		
QY	301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTTACTATATAATAA	360		
DB	1404	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTTACTATATAATAA	1463		
QY	361	GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA	420		
DB	1464	GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTACTATGGTAA	1523		
QY	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGTGGAAATGCCTCAGGCACC	480		
DB	1524	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGTGGAAATGCCTCAGGCACC	1583		
QY	481	AAACCAATATGACCCCTATTTCACATCCAGAAAGCAGCCCAAGACCGCGGAAACTTGGTCTT	540		
DB	1584	AAACCAATATGACCCCTATTTCACATCCAGAAAGCAGCCCAAGACCGCGGAAACTTGGTCTT	1643		
QY	541	ATCTGAATGAAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC	600		
DB	1644	ATCTGAATGAAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC	1703		
QY	601	ACCAATTTACTGTGGACTACAAAGTCTCAATCAGCAAGTAATTTACCTGCTTACATGGA	660		
DB	1704	ACCAATTTACTGTGGACTACAAAGTCTCAATCAGCAAGTAATTTACCTGCTTACATGGA	1763		
QY	661	TAATTAACCTCAAGGAAGTTCATCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTCAC	720		
DB	1764	TAATTAACCTCAAGGAAGTTCATCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTCAC	1823		
QY	721	AACTGGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780		
DB	1824	AACTGGGATGGATGCTTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	1883		
QY	781	TTACATATACAGCAAGTAAGTTGCCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT	840		

DB	1884	TTACAAATACAGACGAATACGTTGCCCTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT	1943
QY	841	TGTTGATGTTTCTAAACGGTAAAGTCATTGCCAGCTAGGACGACGCCATCAGTCAAGTAA	900
DB	1944	TGTTGATGTTTCTAAACGGTAAAGTCATTGCCAGCTAGGACGACGCCATCAGTCAAGTAA	2003
QY	901	TGTTTCTCTCGGAATTAAACCAAGCAGTAGAATAACAAACCGGAGCTGGGGATCAACTATGAA	960
DB	2004	TGTTTCTCTCGGAATTAAACCAAGCAGTAGAATAACAAACCGGAGCTGGGGATCAACTATGAA	2063
QY	961	ACCGATACAGACTATGCTCCTCGCTTGGAGTAGGGTGTCTACGATTCACACTGCTACTAT	1020
DB	2064	ACCGATACAGACTATGCTCCTCGCTTGGAGTAGGGTGTCTACGATTCACACTGCTACTAT	2123
QY	1021	CGTTACAGATGAGCCCTATAACTACCTGGGACAAATFACCTGTTTATTAACATGGGATAG	1080
DB	2124	CGTTACAGATGAGCCCTATAACTACCTGGGACAAATFACCTGTTTATTAACATGGGATAG	2183
QY	1081	GGGCTACTTTGGCAACATCACTTTCGAATACGCCCTGCAACAATTCGGGAACCGTCCGAGC	1140
DB	2184	GGGCTACTTTGGCAACATCACTTTCGAATACGCCCTGCAACAATTCGGGAACCGTCCGAGC	2243
QY	1141	CGTGGAACTCTTAAACAAGTCGAGCTCAACCGGCCCAAGACTTTCTTAATGGTCTAGG	1200
DB	2244	CGTGGAACTCTTAAACAAGTCGAGCTCAACCGGCCCAAGACTTTCTTAATGGTCTCGG	2303
QY	1201	AATCGACTACCCCAAGTATTCACTACTCAATGCAATTCGAATTAACAACAACCGAATCAGA	1260
DB	2304	AATCGACTACCCCAAGTATTCACTACTCAATGCAATTCGAATTAACAACAACCGAATCAGA	2363
QY	1261	CAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGTGCTTACGCTGCTTTTGCAGAAATGG	1320
DB	2364	CAAAAAATATGGAGCAAGTAGTGAAAGATGGCTGTGCTTACGCTGCTTTTGCAGAAATGG	2423
QY	1321	TGGAATCTACTATAAACCAATGATATCCATAAAGTCGCTTTTAGTCATGGGAGTGAATA	1380
DB	2424	TGGAATCTACTATAAACCAATGATATCCATAAAGTCGCTTTTAGTCATGGGAGTGAATA	2483
QY	1381	AGAGTTCTCTTAATGTGCGAACTCGTCCATGAAGAAACACAGACCTATATGATGACCGA	1440
DB	2484	AGAGTTCTCTTAATGTGCGAACTCGTCCATGAAGAAACACAGACCTATATGATGACCGA	2543
QY	1441	CATGATGAAACAGTCTTGTGACTTATGGAAGTGCAGCAATGCTCTTCTGCTTGGCTCCC	1500
DB	2544	CATGATGAAACAGTCTTGTGACTTATGGAAGTGCAGCAATGCTCTTCTGCTTGGCTCCC	2603
QY	1501	TCAGGCTGGTAAACAGGAACTCTAACTATACAGACGAGGAAATTCAAACACCATCAA	1560
DB	2604	TCAGGCTGGTAAACAGGAACTCTAACTATACAGACGAGGAAATTCAAACACCATCAA	2663
QY	1561	GACCTCTCAATTTGTAGCACCTGATGAACATTTTGTGCTGCTATACGCGTAAATTTCAAT	1620
DB	2664	GACCTCTCAATTTGTAGCACCTGATGAACATTTTGTGCTGCTATACGCGTAAATTTCAAT	2723
QY	1621	GGCTGATGGACAGGCTATTTCTAACCGCTGTGACCACTTGTAGGCAATGGCCTTACGGT	1680
DB	2724	GGCTGATGGACAGGCTATTTCTAACCGCTGTGACCACTTGTAGGCAATGGCCTTACGGT	2783
QY	1681	CGCTGCCAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGGAGCAATCCAGAGA	1740
DB	2784	CGCTGCCAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGGAGCAATCCAGAGA	2843
QY	1741	TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATTTTAAAAATGGTGTCTG	1800
DB	2844	TTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGATTTTAAAAATGGTGTCTG	2903
QY	1801	TTCTACGTGGAAGTCACTGCTCCACAACACCCCATCAACTGAAAGTTTCAAGCTCATC	1860
DB	2904	TTCTACGTGGAAGTCACTGCTCCACAACACCCCATCAACTGAAAGTTTCAAGCTCATC	2963
QY	1861	ATCAGATAGTTCAACTTCACAGTCTAGCTCAACACTCCCAAGCACAATTAATAGTAGGAC	1920
DB	2964	ATCAGATAGTTCAACTTCACAGTCTAGCTCAACACTCCCAAGCACAATTAATAGTAGGAC	3023

[illegible]

```

CDS
/ gene="spr0334"
5514..6908
/ gene="spr0334"
/codon_start=1
/transl_table=11
/label=spr0334
/product="Conserved hypothetical protein"
/protein_id="AAK99138.1"
/db_xref="GI:15457891"
/translation="MSKKRRNKKKEAQEPQDFDEAKELTVGQAIRKNEEVEAGVLP
ESLIDKIVKQHRDEADKFAIRYKKEEFVETSLDDILOEMREAEVKSASSEEV
PSEDLILPLDEDEQGLDPLLDENPTETMERVEEQNLSRLDDEDSEKSKKGF
ILTVLALSVITCVAYVYVQVARSKEIETQSQTANQSDVDDENTLYDAFYNSN
KTKLNSOPDKLSLTKLIDKLEGSREHTLAKSKYDSLATQIKAQDYNQAOPEKPAIV
DGLVLTNAKASDAKFTDIKTGNTLKDLDKAIKSLGKSOQTSSSSSSSOFSSSSS
QASSNTSEPKPSSNETHRSSEVNMGLSSAGVAVORSARVAYNQSALDDNSAM
FRADGVQLQILATSRSGYITGDQILERNVINGNYLNKPKDGTLYLFTLNCKTGY
FVNGAGHADDUDY"
6984..8408
/ gene="gnd"
/ note="spr0335"
6984..8408
/ gene="gnd"
/ EC_number="1.1.1.44"
/codon_start=1
/transl_table=11
/label=spr0335
/product="6-phosphogluconate dehydrogenase"
/protein_id="AAK99139.1"
/db_xref="GI:15457892"
/translation="MTKANFGVVGMAVGRNLALNIESRGYTVAIYNSRKEKTEVDIA
CPEKNFVSDVESFVNSIEKPRIMLVQAGPGTDATIQALLPHLDKGDILIDGNN
TPYKDTIRNEELANSINFGTGVSGGKALEGSPIMPGQKEAYELIVADVLEBIS
AKAPEDGKPCVTYIGPDGAGHYVKNVHNGIEYGMOLIAESYDLMOHLLGLSAEDMAE
IFTEKNGKGLSYLLEITADILSRKDDGQDGPVIDYILDAAGNKGTCKWTSQSLDL
GVPLSLITSEVPARYISTYKEERVHASKVLPKPAFNEFGDKAELEIKRQALYSKI
ISYAGQALRVASKENNLPFADIASIWRDGCIIIRSFQKRIITDAYNRDADLANLL
LDEYFDYTKYQQAQRDIVALAVQGVPTFSAITYFDSYRSADLPANLIQAORD
8420..9109
/ gene="csr"
/ note="spr0336"
8420..9109
/ gene="csr"
/ note="Putative"
/codon_start=1
/transl_table=11
/label=spr0336
/product="Response regulator"
/protein_id="AAK99140.1"
/db_xref="GI:15457893"
/translation="MGKRILLLEKERNLAHFLSLELQEQYRVDLVEBQKALSMAIQL
TPDYILLNVNGLMDMAODFAEKLKSRTPKPSVIMILDHWDLEQEEVYVQREAVSYIY
KPVLIENLVARISALFRGRDFIDQHCSLMKVPTVRLNLDVHEHTVYRGEEMIALTR
REYDILLATLNGSKKVLTRQLLESVWKYESATETNIVDVIYIRLSKLDVKGQKSYIK
TVRGVGYTMOE"
9207..10223
/ gene="cbpf"
/ note="spr0337"
9207..10223
/ gene="cbpf"
/codon_start=1
/transl_table=11
/label=spr0337
/product="Choline-binding protein F"
/protein_id="AAK99141.1"

Query Match 71.9%; Score 1438; DB 1; Length 10313;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 198; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATAAATACTCACTGCTGACTTGGGTTCTGAAGCGCGGTCAA 60
|||||

```

```

Db 2115 TAAATCTAGCAATAAATACTCACTGCTGACTTGGGTTCTGAAGCGCGGTCAA 2056
QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCAAGACCA 120
|||||
Db 2055 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCAAGACCA 1996
|||||
QY 121 TCGCTTCTTCGACCACAGGGGGATTGATACCATCCGCTATCCTGGGAGCTTCTTTCGGCAA 180
|||||
Db 1995 TCGCTTCTTCGACCACAGGGGGATTGATACCATCCGCTATCCTGGGAGCTTCTTTCGGCAA 1936
|||||
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCTACCCCAACAGTTGATTAAGTTGAC 240
|||||
Db 1935 TCTGCAAGCAATTCCTCCCAAGGTGATCAACTCTACCCCAACAGTTGATTAAGTTGAC 1876
|||||
QY 241 TTACTTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
|||||
Db 1875 TTACTTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 1816
|||||
QY 301 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
|||||
Db 1815 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 1756
|||||
QY 361 GGTCTACATGCTTAATGGGNACTATGGAATGCAGACAGAGCTCAAAACTACTATGTTAA 420
|||||
Db 1755 GGTCTACATGCTTAATGGGNACTATGGAATGCAGACAGAGCTCAAAACTACTATGTTAA 1696
|||||
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGCCTGGCTGGAATGCCTCAGGCACC 480
|||||
Db 1695 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGCCTGGCTGGAATGCCTCAGGCACC 1636
|||||
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGAGCCCGCGGAAACTTGTCTT 540
|||||
Db 1635 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCCAAGAGCCCGCGGAAACTTGTCTT 1576
|||||
QY 541 ATCTGAATGAAAAAATAAGGCTACACTCTGCTGCAACAGATATGAGAAAGCAGTCAATAC 600
|||||
Db 1575 ATCTGAATGAAAAAATAAGGCTACACTCTGCTGCAACAGATATGAGAAAGCAGTCAATAC 1516
|||||
QY 601 ACCAATATCTGATGGACTACAAAGTCTCAAAAGTCAAAAGTCAAAAGTCAAAAGTCAAAAGT 660
|||||
Db 1515 ACCAATATCTGATGGCTACAAAGTCTCAAAAGTCAAAAGTCAAAAGTCAAAAGTCAAAAGT 1456
|||||
QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAGAAACAGGCTATTAACCTACTCAC 720
|||||
Db 1455 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAGAAACAGGCTATTAACCTACTCAC 1396
|||||
QY 721 AACTGGGATGGATGCTCACAATAATGTAGACCAAGAGCTCAAAAGTCAAAAGTCAAAAGT 780
|||||
Db 1395 AACTGGGATGGATGCTCACAATAATGTAGACCAAGAGCTCAAAAGTCAAAAGTCAAAAGT 1336
|||||
QY 781 TTACAATACAGCAATACGTTGCCCTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT 840
|||||
Db 1335 TTACAATACAGCAATACGTTGCCCTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT 1276
|||||
QY 841 TGTGTATGTTTCTAAGCGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 900
|||||
Db 1275 TGTGTATGTTTCTAAGCGTAAAGTCAATTCGCCAGCTAGGAGCAGCCATCAGTCAAGTAA 1216
|||||
QY 901 TGTGTCTCTCGGAATTTAAACCAAGCAGTAGAACAACCCGCGACTGGGGATCAACTATGAA 960
|||||
Db 1215 TGTGTCTCTCGGAATTTAAACCAAGCAGTAGAACAACCCGCGACTGGGGATCAACTATGAA 1156
|||||
QY 961 ACCGATCACAGACTATGCTCCTGCCCTGGAGTACGGTGTCTACGATTCAACTGCTACTAT 1020
|||||
Db 1155 ACCGATCACAGACTATGCTCCTGCCCTGGAGTACGGTGTCTACGAGTCAACTGCCACTAT 1096
|||||
QY 1021 CGTTCAGATGAGCCCTATAACTACCTGGGACAAATACTCCTGTTTATTAACCTGGGATAG 1080
|||||
Db 1095 CGTTCAGATGAGCCCTATAACTACCTGGGACAAATACCCTGTTTATTAACCTGGGATAG 1036
|||||
QY 1081 GGCGTACTTTGGCAACATCACCTTGAATACGCCCTGCAACATCGGAAACCTGCCAGC 1140
|||||
Db 1035 GGCGTACTTTGGCAACATCACCTTGAATACGCCCTGCAACATCGGAAACCTGCCAGC 976
|||||

```

Qy	1141	CGTGAAGACTCTAAACAAGAGTCGGACCTCAACCGCGCCAGACATTTCTCTAAATGGTCTAGG	1200
Db	975		
Qy	1201	AATCGACTACCAAGTATTCACTACTCAATGCCATTTCAAGTACACAAACCGGAATCAGA	1260
Db	915		
Qy	1261	CAAAAATATGGAGCAAGTAGTGAAGAGTGGCTGCTCTTACGCTGGCTTTGCAAAATGG	1320
Db	855	CAAAAATATGGAGCAAGTAGTGAAGAGTGGCTGCTCTTACGCTGGCTTTGCAAAATGG	796
Qy	1321	TGGAACTTACTATAAACCAATGTATATCCATAAAGTGGTCTTTAGTGATGGAGTGAATA	1380
Db	795	TGGAACCTTACTATAAACCAATGTATATCCATAAAGTGGTCTTTAGTGATGGAGTGAATA	736
Qy	1381	AGAGTTCTCTAATGTGCGAACTCGTGCCATCAAGGAACGACAGCTATATGATGACCGA	1440
Db	735	AGAGTTCTCTAATGTGCGAACTCGTGCCATCAAGGAACGACAGCTATATGATGACCGA	676
Qy	1441	CATGATGAACACAGCTTTGACTTTATGGAACCTGGACGAATGCCATCTTGGCTGCC	1500
Db	675	CATGATGAACACAGCTTTGACTTTATGGAACCTGGACGAATGCCATCTTGGCTGCC	616
Qy	1501	TCAGGCTGGTAAACACAGGAACCTCTAATACTACAGACGAGGAATGGAACACCATCAA	1560
Db	615	TCAGGCTGGTAAACACAGGAACCTCTAATACTACAGACGAGGAATGGAACACCATCAA	556
Qy	1561	GACCTCTCAATTTGTAGCACCTGATGAACATTTTGGCTGCTATACGCGTAAATATTCAAT	1620
Db	555	GACCTCTCAATTTGTAGCACCTGATGAACATTTTGGCTGCTATACGCGTAAATATTCAAT	496
Qy	1621	GGCTGTATGGACAGGCTATTCTACCGTCTGACACCACCTTGTAGGCAATGCCCTTACGGT	1680
Db	495	GGCTGTATGGACAGGCTATTCTACCGTCTGACACCACCTTGTAGGCAATGCCCTTACGGT	436
Qy	1681	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTCAAGGAAGCAATCCAGAAGA	1740
Db	435	CGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGTCTCAAGGAAGCAATCCAGAAGA	376
Qy	1741	TTGGAATATACGAGAGGGCTCTACAGAAATGGAGAATTCGTATTTAAAATGGTGCTCG	1800
Db	375	TTGGAATATACGAGAGGGCTCTACAGAAATGGAGAATTCGTATTTAAAATGGTGCTCG	316
Qy	1801	TTCTACGTGGAACCTCACCTGCTCCACAAACACCCCATCAACTGAAAGTTTCAAGCTCATC	1860
Db	315	TTCTACGTGGAACCTCACCTGCTCCACAAACACCCCATCAACTGAAAGTTTCAAGCTCATC	256
Qy	1861	ATCAGATAGTTTCAACTTTCACAGTCTAGTCAACCACTCCCAAGCACAATAATAGTACGAC	1920
Db	255	ATCAGATAGTTTCAACTTTCACAGTCTAGTCAACCACTCCCAAGCACAATAATAGTACGAC	196
Qy	1921	TACCAATCCTTAACAATAATACGCAACAATCAAAATACACCCCTCATCAACAAAATCAGAA	1980
Db	195	TACCAATCCTTAACAATAATACGCAACAATCAAAATACACCCCTCATCAACAAAATCAGAA	136
Qy	1981	TCCTCAACCCAGCAACAACCA	1999
Db	135	TCCTCAACCCAGCAACAACCA	117

RESULT	11
AX111327	
LOCUS	AX111327 1249 bp DNA linear PAT 30-APR-2001
DEFINITION	Sequence 2060 from Patent WO0123604.
ACCESSION	AX111327
VERSION	AX111327.1 GI:13927619
KEYWORDS	.
SOURCE	Streptococcus pneumoniae.
ORGANISM	Streptococcus pneumoniae Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.

REFERENCE	1 (bases 1 to 1249)
AUTHORS	Bergeron,M.G., Boisjolinot,M., Huletsky,A., m Nard,C., Ouellette,M., Picard,F.J. and Roy,P.H.
TITLE	Highly conserved genes and their use to generate probes and primers for detection of microorganisms
JOURNAL	Patent: WO 0123604-A 2060 05-APR-2001;
FEATURES	Infectio Diagnostic (I.D.I.) INC. (CA)
source	Location/Qualifiers 1. .1249 /organism="Streptococcus pneumoniae" /strain="StrR-57" /db_xref="taxon:1313"
BASE COUNT	402 a 302 c 246 g 299 t
ORIGIN	
Query Match	59.9%; Score 1198; DB 6; Length 1249;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 1248; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	720 CAACGTGGATGGATGCTACACAAATGTAGACCAAGAAGCTCAAAAACATCTGTGGGATA 779
Db	1 CAACGTGGATGGATGCTACACAAATGTAGACCAAGAAGCTCAAAAACATCTGTGGGATA 60
Qy	780 TTTACAATACAGACGAATACGTTGCTATCCAGACGATGAATGCAAGTCGCTTCTACCA 839
Db	61 TTTACAATACAGACGAATACGTTGCTATCCAGACGATGAATGCAAGTCGCTTCTACCA 120
Qy	840 TTGTTGATGTTTCTAACGGTAAAGTCATGCCCCAGCTAGGACGCGCATCAGTCAAGTA 899
Db	121 TTGTTGATGTTTCTAACGGTAAAGTCATGCCCCAGCTAGGACGCGCATCAGTCAAGTA 180
Qy	900 ATGTTTCCTTCGGAAATTACCAAGCAGTAGAAAACAAACCGGACCTGGGGATCAACTATGA 959
Db	181 ATGTTTCCTTCGGAAATTACCAAGCAGTAGAAAACAAACCGGACCTGGGGATCAACTATGA 240
Qy	960 AACCGATCACAGACTATGCTCCTGCTTGGAGTAGCGTGTCTACGATTCAACTGCTACTA 1019
Db	241 AACCGATCACAGACTATGCTCCTGCTTGGAGTAGCGTGTCTACGATTCAACTGCTACTA 300
Qy	1020 TCGTTACAGTAGGCCCTATAACTACCCCTGGACAAATACCTCTGTTTATAACTGGGATA 1079
Db	301 TCGTTACAGTAGGCCCTATAACTACCCCTGGACAAATACCTCTGTTTATAACTGGGATA 360
Qy	1080 GGGCTACTTTGGCAACATCATCTTGGCAATACGCCCTGCAACATCGCGAACGTCGCCAG 1139
Db	361 GGGCTACTTTGGCAACATCATCTTGGCAATACGCCCTGCAACATCGCGAACGTCGCCAG 420
Qy	1140 CCGTGGAAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAG 1199
Db	421 CCGTGGAAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTCCTAAATGGTCTAG 480
Qy	1200 GAATCGACTACCCAGTATTCACTACTCAATGCCATTTTCAAGTAACACACACCGAATCAG 1259
Db	481 GAATCGACTACCCAGTATTCACTACTCAATGCCATTTTCAAGTAACACACACCGAATCAG 540
Qy	1260 ACAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGTGCTTACGCTGCCCTTTGCAAAATG 1319
Db	541 ACAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGTGCTTACGCTGCCCTTTGCAAAATG 600
Qy	1320 GTGGAACCTTACTATAAACCAATGTATATCCATAAGTCGTCTTTAGTGATGGGAGTGAAA 1379
Db	601 GTGGAACCTTACTATAAACCAATGTATATCCATAAAGTCGTCTTTAGTGATGGGAGTGAAA 660
Qy	1380 AAGAGTTCTCTAATGTTCGGAACCTCGTGCCATGAAGAAACGACAGCCTATATGATGACCG 1439
Db	661 AAGAGTTCTCTAATGTTCGGAACCTCGTGCCATGAAGAAACGACAGCCTATATGATGACCG 720
Qy	1440 ACATGATCAAAACAGTCTTGACTTATGGAACCTGGACGAAATGCCTATCTCTTGGCTCC 1499
Db	721 ACATGATCAAAACAGTCTTGACTTATGGAACCTGGACGAAATGCCTATCTCTTGGCTCC 780
Qy	1500 CTCAGGCTGGTAAAAACAGGAACCTCTAACTATACAGACGAGGAAATTTGAAAACCACTCA 1559


```
|||||
Db 781 CTCAGGCTGGTAAACAGAACTCTAACTATACAGACGAGGAAATGAAAACCATCA 840
QY 1560 AGACCTCTCAATTTGAGCACCTGATGAACCTATTTGCTGGCTATACGCTTAATATTCAA 1619
Db 841 AGACCTCTCAATTTGAGCACCTGATGAACCTATTTGCTGGCTATACGCTTAATATTCAA 900
QY 1620 TGGCTGTATGGACAGGCTATTTCTAACCGTCTGACACCACTTTGAGCAATGGCCTTACGG 1679
Db 901 TGGCTGTATGGACAGGCTATTTCTAACCGTCTGACACCACTTTGAGCAATGGCCTTACGG 960
QY 1680 TGGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGGAACAATFCCAGAAG 1739
Db 961 TGGCTGCCAAAGTTTACCGCTCTATGATGACCTACCTGCTCTGAAGGAACAATFCCAGAAG 1020
QY 1740 ATTGGAATATACAGAGGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGGTGCTC 1799
Db 1021 ATTGGAATATACAGAGGGGCTCTACAGAAATGGAGAATTCGTATTTAAAAATGGTGCTC 1080
QY 1800 GTTCTAGTGGAACTCACCTGCTCCCAACAACCCCACTCAACTGAAAGTTCAAGCTCAT 1859
Db 1081 GTTCTAGTGGAGCTCACCTGCTCCCAACAACCCCACTCAACTGAAAGTTCAAGCTCAT 1140
QY 1860 CATCAGATAGTCAACTTACAGTCTAGCTCAACCACTCCAGCAGCACAATAATAGTACGA 1919
Db 1141 CATCAGATAGTCAACTTACAGTCTAGCTCAACCACTCCAGCAGCACAATAATAGTACGA 1200
QY 1920 CTACCAATCCTTACCAATAATACGCAACAATCAAAATCAACCCCTGATCA 1968
Db 1201 CTACCAATCCTTACCAATAATACGCAACAATCAAAATCAACCCCTGATCA 1249

RESULT 12
AB006874
LOCUS      AB006874               1195 bp      DNA      linear      BCT 13-FEB-1999
DEFINITION Streptococcus pneumoniae pbpla gene for penicillin binding protein
            1A, partial cds, isolate:#10/219.
ACCESSION AB006874
VERSION   AB006874.1  GI:3395651
KEYWORDS  pbpla; penicillin binding protein 1A.
SOURCE    Streptococcus pneumoniae (isolate:#10/219) DNA.
ORGANISM  Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE  1 (bases 1 to 1195)
AUTHORS   Asahi,Y.
DIRECT SUBMISSION
TITLE      Submitted (27-AUG-1997) Yasuko Asahi, Teikyo University School of
JOURNAL    Medicine, Department of Clinical Pathology; 2-11-1 Kaga,
            Itabashi-ku, Tokyo 173-8605, Japan
            (E-mail: asahi@med.teikyo-u.ac.jp, Tel:81-3-3964-1211,
            Fax:81-3-3963-6023)
REFERENCE  2 (sites)
AUTHORS   Asahi,Y. and Ubukata,K.
TITLE      Association of a thr-371 substitution in a conserved amino acid
            motif of penicillin-binding protein 1A with penicillin resistance
            of Streptococcus pneumoniae
JOURNAL    Antimicrob. Agents Chemother. 42 (9), 2267-2273 (1998)
MEDLINE    98409715
FEATURES   Location/Qualifiers
            1..1195
            /organism="Streptococcus pneumoniae"
            /isolate="#10/219"
            /db_xref="taxon:1313"
            1..1195
            /gene="pbpla"
            <1..>1195
            /gene="pbpla"
            /codon_start=1
            /transl_table=11
            /product="penicillin binding protein 1A"
            /protein_id="BAA32070.1"
            /db_xref="GI:3395652"
```

```
/translation="VAYPDELOQVASTIVDSNGKVIQALGARHOSSNVSFGINQAVE
TNRDWSMKPITDYAPALEYGVYDSTATIVHDEPNYNGTNTPVYNNDRGVFGNITL
QYALQOSNRNPVAVETLNKVLNRKTEFLNGLIDYPSIHYSNAINSTTMSDKKYGAS
SEKMAAAYAAAFANGCTYKPKYIHKVFSDDGSEKESNVGTTRAMKETAYMTDMKMT
VLUSYGTNRNLAYLAWLPQAGKVTSNYTDDEIENHIIKTSQFVAPDFAGYTRKYSWAV
WTGYSNRLTPLVGNGLTAAKVRSMYTLSEGSNPEDWNIPEGLYRNGEYFKNGAR
STWNSPAPQPPSPSTESSSSDSSTSQSSSTTPTNNSTTDPNNNTQOOSNTTPDOON
ONPOPA"
BASE COUNT      381 a      297 c      233 g      284 t
ORIGIN
Query Match      44.5%; Score 889; DB 1; Length 1195;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 800 GTTGCCCTATCCAGACGATGAATTCGAAGTCGCTTCTACCATGTTGTTGTTTACCGGT 859
Db 1 GTTGCCCTATCCAGACGATGAATTCGAAGTCGCTTCTACCATGTTGTTGTTTACCGGT.60
QY 860 AAGTCAATTCGCCAGCTAGGAGCAGCCCATCAGTCAAGTAATGTTTCTTCGGAATTAAC 919
Db 61 AAGTCAATTCGCCAGCTAGGAGCAGCCCATCAGTCAAGTAATGTTTCTTCGGAATTAAC 120
QY 920 CAAGCAGTAGAACAACAAACCCGACTGGGATCAACTATGAAACCGATCAGACTATGCT 979
Db 121 CAAGCAGTAGAACAACAAACCCGACTGGGATCAACTATGAAACCGATCAGACTATGCT 180
QY 980 CTTGCCCTTGGAGTACGGTGTCTACGATTCAACTGCTACTATCGTTCAAGTACGCCCTAT 1039
Db 181 CTTGCCCTTGGAGTACGGTGTCTACGATTCAACTGCTACTATCGTTCAAGTACGCCCTAT 240
QY 1040 AACTACCCCTGGGACAATACTCTCTGTTTAACTGGGATAGGGGCTACTTTTGCAACATC 1099
Db 241 AACTACCCCTGGGACAATACTCTCTGTTTAACTGGGATAGGGGCTACTTTTGCAACATC 300
QY 1100 ACCTTGCAATACGCCCTCGCAACAATCGAAGACGTCGCCAGCCGTGGAACACTCTAAACAAG 1159
Db 301 ACCTTGCAATACGCCCTCGCAACAATCGAAGACGTCGCCAGCCGTGGAACACTCTAAACAAG 360
QY 1160 GTGCGGACTCAACGGCGCCCAAGACTTTCTTAATGGTCTAGGAATCGACTACCCAAGTATT 1219
Db 361 GTGCGGACTCAACGGCGCCCAAGACTTTCTTAATGGTCTAGGAATCGACTACCCAAGTATT 420
QY 1220 CACTACTCAAAATGCCATTTCAAGTAACAACAACCGAATCAGACAAAAAATATGGAGCAAGT 1279
Db 421 CACTACTCAAAATGCCATTTCAAGTAACAACAACCGAATCAGACAAAAAATATGGAGCAAGT 480
QY 1280 AGTGAAGAAGATGGCTGCTGCTTACGCTGCTTTGCAAAATGGTGGAACTTACTATAACCA 1339
Db 481 AGTGAAGAAGATGGCTGCTGCTTACGCTGCTTTGCAAAATGGTGGAACTTACTATAACCA 540
QY 1340 ATGTATATCCATAAAGTCGCTTTTAGTGTGGAGTGAAAGAGGTCTCTAATGTGCGA 1399
Db 541 ATGTATATCCATAAAGTCGCTTTTAGTGTGGAGTGAAAGAGGTCTCTAATGTGCGA 600
QY 1400 ACTCGTGCCATGAAGGAAACGACAGCCTATATGATCACCAGCATGATGAAACAGTCTTG 1459
Db 601 ACTCGTGCCATGAAGGAAACGACAGCCTATATGATCACCAGCATGATGAAACAGTCTTG 660
QY 1460 ACTTATGGAACGGGAGAAATGCCTATCTTGTGCTGCCCTCCCTCAGGCTGATAAACAGGA 1519
Db 661 AGTTATGGAACGGGAGAAATGCCTATCTTGTGCTGCCCTCCCTCAGGCTGATAAACAGGA 720
QY 1520 ACCTCTAACTATACAGCAGGAGAAATTTGAAACCAACATCAAGACCTCTCAATTTGTAGCA 1579
Db 721 ACCTCTAACTATACAGCAGGAGAAATTTGAAACCAACATCAAGACCTCTCAATTTGTAGCA 780
QY 1580 CTTGTGAACTATTGCTGGCTATACGCGTAAATATTTCAATGGCTCTATGGACAGGCTAT 1639
Db 781 CTTGACGAACTATTGCTGGCTATACGCGTAAATATTTCAATGGCTCTATGGACAGGCTAT 840
QY 1640 TCTAACCGCTGTGACACCACTTGTAGGCAATGGCCTTACGGTGCCTGCCAAGATTTACCGC 1699
```



```
|||||
Db 964 TACAGAAATGGAGATTCGTATTTAAAAATGGTGCCTTACGTGGAGCTCACGTGCT 1023
QY 1823 CCACAAACACCCCACTCAACTGAAAGTTCAAGCTCATCATCATGATAGTTCAACTTCCACAG 1882
Db 1024 CCACAAACACCCCACTCAACTGAAAGTTCAAGCTCATCATCATGATAGTTCAACTTCCACAG 1083
QY 1883 TCTAGCTCAACCACTCCAAAGCACAAATAATAGTACGACTACCAATCCTTAACAATAATAGC 1942
Db 1084 TCTAGCTCAACCACTCCAAAGCACAAATAATAGTACGACTACCAATCCTTAACAATAATAGC 1143
QY 1943 CAACAAATCAAAATACAAACCCCTGATCAACAAATAATCAGAATCTTCAACCAAGCAC 1994
Db 1144 CAACAAATCAAAATACAAACCCCTGATCAACAAATAATCAGAATCTTCAACCAAGCAC 1195

RESULT 14
AR084885
LOCUS AR084885 960 bp DNA linear PAT 01-SEP-2000
DEFINITION Sequence 3 from patent US 5981229.
ACCESSION AR084885
VERSION AR084885.1 GI:10011656
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 960)
AUTHORS Masure H, Robert, Pearce B.J. and Tuomanen, E.
TITLE Bacterial exported proteins and acellular vaccines based thereon
JOURNAL Patent: US 5981229-A 3 09-NOV-1999;
FEATURES
    Location/Qualifiers
        1..960
            /organism="unknown"
BASE COUNT 315 a 224 c 182 g 239 t
ORIGIN
```

```
Query Match 39.8%; Score 796; DB 6; Length 960;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 TAAATCTACGACATAAATAAATCAACTCATTCGCTGACTTGGGTTCTGAACGGCGGTCAA 60
Db 12 TAAATCTACGACATAAATAAATCAACTCATTCGCTGACTTGGGTTCTGAACGGCGGTCAA 71
QY 61 TCGCCAAAGCTAATGATATTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 72 TCGCCAAAGCTAATGATATTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 131
QY 121 TCGCTTCTTCGACCACAGGGGATTGATACCATCCGTATCCTGGGAGCTTCTTGGCGAA 180
Db 132 TCGCTTCTTCGACCACAGGGGATTGATACCATCCGTATCCTGGGAGCTTCTTGGCGAA 191
QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCACTCTCACCACAGAGTTGATTAAGTTGAC 240
Db 192 TCTGCAAGCAATTCCTCCAGGTGGATCACTCTCACCACAGAGTTGATTAAGTTGAC 251
QY 241 TTACTTTTCAACTTCGACTTCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT 300
Db 252 TTACTTTTCAACTTCGACTTCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT 311
QY 301 ACGGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 360
Db 312 ACGGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAATAA 371
QY 361 GGTCTACATGCTTAATGGGAATCTGGAATGCAGACAGAGCTCAAACTACTATGGTAA 420
Db 372 GGTCTACATGCTTAATGGGAATCTGGAATGCAGACAGAGCTCAAACTACTATGGTAA 431
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGCTGGAATCCCTCAGGCACC 480
Db 432 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGCTGGAATCCCTCAGGCACC 491
```

```
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACCGCGGAACTTGGTCTT 540
Db 492 AAACCAATATGACCCCTATTTCACATCCAGAAGCAGCCCAAGACCGCGGAACTTGGTCTT 551
QY 541 ATCTGAATGAAATCAAGGCTAGATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 552 ATCTGAATGAAATCAAGGCTAGATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 611
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAATTTACCTCTGTACATGGA 660
Db 612 ACCAATTACTGATGGCTACAAGTCTCAAAATCAGCAAGTAATTTACCTCTGTACATGGA 671
QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGAAACAGGCTATAAACCCTACTCAC 720
Db 672 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAAGAAACAGGCTATAAACCCTACTCAC 731
QY 721 AACTGGGATGGATGCTTACACAAATGTAGACCAAGGCTCAAAACATCTGTGGGATAT 780
Db 732 AACTGGGATGGATGCTTACACAAATGTAGACCAAGGCTCAAAACATCTGTGGGATAT 791
QY 781 TTACAAATACAGAGGATAGTTCCTCTATCCAGACGATGAATTCGAAGTGCCTTCTTACCAT 840
Db 792 TTACAAATACAGAGGATAGTTCCTCTATCCAGACGATGAATTCGAAGTGCCTTCTTACCAT 851
QY 841 TGTTCATGTTTCTACGGTAAAGTCAATGCCAGCTAGGACGACGCGCATCAGTCAAGTAA 900
Db 852 TGTTCATGTTTCTACGGTAAAGTCAATGCCAGCTAGGACGACGCGCATCAGTCAAGTAA 911
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTATAAACAACCGCGGACTGGGA 949
Db 912 TGTTCCTTCGGAATTAACCAAGCAGTATAAACAACCGCGGACTGGGA 960

RESULT 15
SPON6
LOCUS S.pneumoniae (8250) ponA gene for penicillin-binding protein 1a.
DEFINITION x67871
ACCESSION x67871.1 GI:47415
VERSION
KEYWORDS penicillin-binding protein 1a; ponA gene.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae.
REFERENCE 1 (bases 1 to 1826)
AUTHORS Hakenbeck, R.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek. Genetik, Ihnestr. 73, 1000 Berlin 33, FRG
AUTHORS Martin, C., Sibold, C. and Hakenbeck, R.
TITLE Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain
JOURNAL EMBO J. 11 (11), 3831-3836 (1992)
MEDLINE
FEATURES
    Location/Qualifiers
        1..1826
            /organism="Streptococcus pneumoniae"
            /strain="8250"
            /db_xref="taxon:1313"
            /clone="pcMK1-pcMK16"
            1..1826
                /gene="ponA"
                1..>1826
                    /gene="ponA"
                    /codon_start=1
                    /transl_table=11
                    /product="penicillin-binding protein 1a"
                    /protein_id="CAA48071.1"
                    /db_xref="GI:47415"
                    /db_xref="SPTREMBL:Q54951"
                    /translation="MNKOTILRIAKYVSICTFLTITAVMLGGGIFLYVVKAPALSE"
    source
```

SKLVATTSKIVDKNOLLIADLGSERRVNAQANDIPTDLVKAIVSIEDHFFDHRGID
TIRILGATLRNLSQSLQSGSTLTOOLIKLYTSTSDQTSRKAQEAWLAIQLEQK
ATKQILTYIYINKVYMSNGYGMOTAAQNYGKDLNLSLPOLALLAGPQAPNOYDP
YSHPEAQRRLNLVSEMKNOGYISAEQYKAVNTPITDGLQSLKASANYPAYMDNYL
KEVINQVEETCYNLLTTCMDVYTVNQDEAKHLWDIYNTDEVVAYPDELOVASTIV
DYSNGKVIQAIGARHQSSNVSGINQAVETNRDWSGTMKPIIDYAPALEYGYVDSTAT
IVHDEPYNPYGTDTPIYNWDRGFYFNGITLQYALQOSNRNPAVETLNKVLGNLRAKFTFLN
GLGIDYPSLHYSNAISSNTTSDSKYIGASSERKMAAAYAAFANGYFYKPMYIHKVYFVS
DGSEKFEFNVGTFRAMKEATAYMTEMKMTVLTYGICRGAYLFWLPQAGKTGTSNTYDE
ETIEKYIKNTGYVADPEMFVGYTRKYAMAVWVTGYSNRLTPIIG"
BASE COUNT 582 a 422 c 364 g 458 t
ORIGIN

Query Match 38.7%; Score 774; DB 1; Length 1826;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 924; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATAAATACTCAATTCGTGACTGGGTCTGAACGCCGCGCTCAA 60
DB 159 TAAATCTACGACATAAATACTCAATTCGTGACTGGGTCTGAACGCCGCGCTCAA 218

QY 61 TGCCCAAGCTAATGATATTCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
DB 219 TGCCCAAGCTAATGATATTCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278

QY 121 TCGCTTTCTCGACACAGGGGATGATACCATCCGTATCTCTGGAGCTTTCTTGCGCAA 180
DB 279 TCGCTTTCTCGACACAGGGGATGATACCATCCGTATCTCTGGAGCTTTCTTGCGCAA 338

QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCACTCTACCCACAGTTGATTAAGTTGAC 240
DB 339 TCTGCAAGCAATTCCTCCAGGTGGATCACTCTACCCACAGTTGATTAAGTTGAC 398

QY 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
DB 399 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458

QY 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTTCGACTACTATATAATAA 360
DB 459 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTTCGACTACTATATAATAA 518

QY 361 GGTCTACATGCTTAATGGGAACATATGGAATGACAGACAGCTCAAACTACTATGTTAA 420
DB 519 GGTCTACATGCTTAATGGGAACATATGGAATGACAGACAGCTCAAACTACTATGTTAA 578

QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGCTGGAATGCTCAGGCACC 480
DB 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGCTGGAATGCTCAGGCACC 638

QY 481 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACGCCGCAAACTTGGCTTT 540
DB 639 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACGCCGCAAACTTGGCTTT 698

QY 541 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGCTCAATAC 600
DB 699 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGCTCAATAC 758

QY 601 ACCAATTACTGTGGTACTACAAGTCTCAAACTCAGCAAGTAATTTACCTGCTTACATGGA 660
DB 759 ACCAATTACTGTGGTACTACAAGTCTCAAACTCAGCAAGTAATTTACCTGCTTACATGGA 818

QY 661 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAGAAACAGGCTATAAACCCTACTCAC 720
DB 819 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAGAAACAGGCTATAAACCCTACTCAC 878

QY 721 AACTGGGATGGATGCTTACAAATGCTAGACCAAGCTCAAAACATCTGTGGGATAT 780
DB 879 AACTGGGATGGATGCTTACAAATGCTAGACCAAGCTCAAAACATCTGTGGGATAT 938

QY 781 TTACATACAGAGATAGTTCCTTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
DB 939 TTACATACAGAGATAGTTCCTTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 998

QY 841 TGTGTGATGTTCTAACGGTAAAGTCATTGCCAGCTAGGACGCGCATCAGTCAAGTAA 900
DB 999 TGTGTGATGTTCTAACGGTAAAGTCATTGCCAGCTAGGACGCGCATCAGTCAAGTAA 1058

QY 901 TGTTTCTCTCGGAATTAACCAAGCAGT 927
DB 1059 TGTTTCTCTCGGAATTAACCAAGCAGT 1085

RESULT 16
AF210748 2154 bp DNA linear BCT 17-SEP-2001
LOCUS Streptococcus pneumoniae isolate SP 1258 penicillin-binding protein
DEFINITION 1A (pbpla) gene, partial cds.
ACCESSION AF210748
VERSION AF210748.1 GI:6563342
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group: Streptococcaceae;
REFERENCE 1 (bases 1 to 2154)
AUTHORS Ferroni, A. and Berche, P.
TITLE Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children
J. Med. Microbiol. 50 (9), 828-832 (2001)
JOURNAL 21432820
MEDLINE 11549185
PUBMED
REFERENCE 2 (bases 1 to 2154)
AUTHORS Ferroni, A. and Berche, P.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1999) Microbiology, Hospital
Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France
FEATURES
source
1..2154
/organism="Streptococcus pneumoniae"
/serotype="23F"
/isolate="SP 1258"
/db_xref="taxon:1313"
1..>2154
/gene="pbpla"
1..>2154
/gene="pbpla"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAF17258.1"
/db_xref="GI:6563343"
/translation="MNKPTILRLIKYLSIFLSLVAIAIVLGGVFFYVSKAPSLSE
SKLVATTSKIVDKNOLLIADLGSERRVNAQANDIPTDLVKAIVSIEDHFFDHRGID
TIRILGATLRNLSQSLQSGSTLTOOLIKLYTSTSDQTSRKAQEAWLAIQLEQK
ATKQILTYIYINKVYMSNGYGMOTAAQNYGKDLNLSLPOLALLAGPQAPNOYDP
YSHPEAQRRLNLVSEMKNOGYISAEQYKAVNTPITDGLQSLKASANYPAYMDNYL
KEVINQVEETCYNLLTTCMDVYTVNQDEAKHLWDIYNTDEVVAYPDELOVASTIV
DYSNGKVIQAIGARHQSSNVSGINQAVETNRDWSGTMKPIIDYAPALEYGYVDSTAT
IVHDEPYNPYGTDTPIYNWDRGFYFNGITLQYALQOSNRNPAVETLNKVLGNLRAKFTFLN
GLGIDYPSLHYSNAISSNTTSDSKYIGASSERKMAAAYAAFANGYFYKPMYIHKVYFVS
DGSEKFEFNVGTFRAMKEATAYMTEMKMTVLTYGICRGAYLFWLPQAGKTGTSNTYDD
ELENTIKNTGYVADPEMFVGYTRKYMAVWVTGYSNRLTPIVGDGFFVAAKYVRSMYTY
LSEDDNNPQDWTPEGLYRSGEFFVKRARSAMTAPAPQAPPTPESSSTSSSTSSSS
STTPTNNNSANNNTNQPNPTTTPGQONQONQNPQ"
BASE COUNT 678 a 497 c 419 g 559 t 1 others
ORIGIN

Query Match 35.3%; Score 705; DB 1; Length 2154;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 855; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATAAATACTCAATTCGTGACTGGGTCTGAACGCCGCGCTCAA 60
|||||

Db 159 TAAATCTACGCAATAAATAAATCAACTCATTTGCTGACTTTGGTGTCTTGAACGGCGCTCAA 218
QY 61 TGCCCAAGCTAATGATATTCCACACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGCCCAAGCTAATGATATTCCACACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTTCTTCGACCACAGGGGATGATACCATCCGTATCCCTGGGAGCTTTCTTTCGGCAA 180
Db 279 TCGCTTTCTTCGACCACAGGGGATGATACCATCCGTATCCCTGGGAGCTTTCTTTCGGCAA 338
QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCACTCTCACCACAGCTTCAATTAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCAGGTGGATCACTCTCACCACAGCTTCAATTAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTTGGTT 300
Db 399 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
Db 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
QY 361 GGTCTACATGCTTAATGGGAATGTAATGCAGACAGCTCAAACTACTATGGTAA 420
Db 519 GGTCTACATGCTTAATGGGAATGTAATGCAGACAGCTCAAACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGAAATGCCCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGAAATGCCCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAGAGCCCAAGACCCGCGGAAACTTTGCTTT 540
Db 639 AAACCAATATGACCCCTATTTCACATCCAGAGAGCCCAAGACCCGCGGAAACTTTGCTTT 698
QY 541 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCAATCAATAC 600
Db 699 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCAATCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAATTTAGCCCTGCTTACATGA 660
Db 759 ACCAATTACTGATGGACTACAAGTCTCAAAATCAGCAAGTAATTTAGCCCTGCTTACATGA 818
QY 661 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAAGCAAGGCTATAAAGCTACTAC 720
Db 819 TAATTACCTCAAGGAAGTCAATCAAGTTGAAGAAAGCAAGGCTATAAAGCTACTAC 878
QY 721 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT 780
Db 879 AACTGGGATGGATGCTACACAAATGTAGACCAAGAGCTCAAAACATCTGTGGGATAT 938
QY 781 TTAAATACAGACGAATAGCTTGCCTTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT 840
Db 939 TTAAATACAGACGAATAGCTTGCCTTATCCAGACGATGAATTCGAAGTCGCTTCTACCAT 998
QY 841 TGTGTGATGTTTCTACGG 858
Db 999 TGTGTGATGTTTCTACGG 1016

RESULT 17
AF046233
LOCUS AF046233 930 bp DNA linear BCT 30-JUN-1998
DEFINITION Streptococcus pneumoniae strain 64429 penicillin-binding protein 1A
(ponA) gene, partial cds.
ACCESSION AF046233
VERSION AF046233.1 GI:2911411
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
REFERENCE 1 (bases 1 to 930)
AUTHORS Smith,A.M. and Klugman,K.P.

TITLE Alterations in PBP 1A essential for high-level penicillin resistance in Streptococcus pneumoniae
JOURNAL Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
MEDLINE 98287565
REFERENCE 2 (bases 1 to 930)
AUTHORS Smith,A.M.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg, Gauteng 2000, South Africa
FEATURES
source Location/Qualifiers
1..930
/organism="Streptococcus pneumoniae"
/strain="64429"
/db_xref="taxon:1313"
<1..>930
/gene="ponA"
<1..>930
/gene="ponA"
/note="PBP 1A"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAC24698.1"
/db_xref="GI:2911412"
/translation="LWDIYNTDEYVAYPDDELQVASTIVDSNGKVIQALGARHOSNNVSFGINQAVETNRDWSMKPTIDYAPALEYGVYDSTATIVHDEPNYPGTNTPVYNNWDRGYFGNITLOVALQSSRNVPAYETLNKVLNRAKTLFNLGIDYPSIHSYNAISNTTESDKKYGASSEKMAAAYAAFAANGFTYKPMYIHKVFSDDGSEKFSNVGTRAMKETTAYMTTMMKTVLVYIGIGRCAYLPWLPQAGKTGTSNYTDEIEIKYIKNTGYVAPDENFVGYTRYAMAVMTGYSNRLTPLVNGLTVAAKVIR"
BASE COUNT 288 a 215 c 198 g 229 t
ORIGIN
Query Match 33.6%; Score 671; DB 1; Length 930;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 770 CTGTGGGATATTACAAATACAGACGAATACGTTGCCCTATCCAGACGATGAATTCGAAGTC 829
Db 1 CTGTGGGATATTACAAATACAGACGAATACGTTGCCCTATCCAGACGATGAATTCGAAGTC 60
QY 830 GCTTTCACCATTTGATGTTCTTAACGGTAAAGTCATTGCCAGCTAGGAGCACGCCAT 889
Db 61 GCTTTCACCATTTGATGTTCTTAACGGTAAAGTCATTGCCAGCTAGGAGCACGCCAT 120
QY 890 CAGTCAAGTAAGTCTTCCTCGGAATTAACCAAGCAGTAGAACAACCCGACTGGGA 949
Db 121 CAGTCAAGTAAGTCTTCCTTCGGAATTAACCAAGCAGTAGAACAACCCGACTGGGA 180
QY 950 TCAACTATGAACCCGATCACAGACTATGCTTCCTCGCTTGGAGTACGGTGTCTACGATTCA 1009
Db 181 TCAACTATGAACCCGATCACAGACTATGCTTCCTCGCTTGGAGTACGGTGTCTACGATTCA 240
QY 1010 ACTGCTACTATGTTCCAGATGAGCCCTATACTACCTGGGCAAAATACTCTCTGTTAT 1069
Db 241 ACTGCTACTATGTTCCAGATGAGCCCTATACTACCTGGGCAAAATACTCTCTGTTAT 300
QY 1070 AACTGGGATAGGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACAATCGGA 1129
Db 301 AACTGGGATAGGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACAATCGGA 360
QY 1130 AACGTCGCCAGCGCTGAAACTCTTAAACAAGTTCGGACTCAACCGCGCAAGACTTTTCCTA 1189
Db 361 AACGTCGCCAGCGCTGAAACTCTTAAACAAGTTCGGACTCAACCGCGCAAGACTTTTCCTA 420
QY 1190 AATGGTCTAGGAATGACATACCCAACTATTCTACTACTCAAAATGCCATTTCAAGTAACACA 1249
Db 421 AATGGTCTAGGAATGACATACCCAACTATTCTACTACTCAAAATGCCATTTCAAGTAACACA 480
QY 1250 ACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTTACGCTGCC 1309
Db 481 ACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTTACGCTGCC 540

```
QY 1310 TTTGCAATGGTGAACCTTACTATAAACAATGTATATCCATAAAGTCGTCTTTAGTGAT 1369
|||||
Db 541 TTTGCAATGGTGAACCTTACTATAAACAATGTATATCCATAAAGTCGTCTTTAGTGAT 600

QY 1370 GGGAGTGAAGAAGAGTTCTCTAATGTCGGAACCTCGTGCATGAAGAAACGACGCCTAT 1429
|||||
Db 601 GGGAGTGAAGAAGAGTTCTCTAATGTCGGAACCTCGTGCATGAAGAAACGACGCCTAT 660

QY 1430 ATGATGACCGA 1440
|||||
Db 661 ATGATGACCGA 671

RESULT 18
AF046235 930 bp DNA linear BCT 30-JUN-1998
LOCUS Streptococcus pneumoniae strain 7851 penicillin-binding protein 1A
DEFINITION (ponA) gene, partial cds.
ACCESSION AF046235
VERSION AF046235.1 GI:2911415
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 930)
AUTHORS Smith,A.M. and Klugman,K.P.
TITLE Alterations in PBP 1A essential for high-level penicillin
resistance in Streptococcus pneumoniae
JOURNAL Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
MEDLINE 98287565
REFERENCE 2 (bases 1 to 930)
AUTHORS Smith,A.M.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg,
Gauteng 2000, South Africa
FEATURES
source
1. .930
Location/Qualifiers
/organism="Streptococcus pneumoniae"
/strain="7851"
/db_xref="taxon:1313"
<1..>930
/gene="ponA"
<1..>930
/gene="ponA"
/feature="PBP 1A"
/codon_start=1
/transl_table=1
/product="penicillin-binding protein 1A"
/protein_id="AAC24700.1"
/db_xref="GI:2911416"
/transl_table="1"
/translation="LMDIYNTDEVYVYPDDELQVASTIVDSNGKVIAQLGARHQSSN
VSGINQAVETNRDWSGTMKPTIDYFALEYGVYDSTATIVHDEPNYPTGNTPVYNN
DRGYGNTLOYLQQRNPVAVETLNKVLNRAKTLFNLGLIDIPDPSIHYSNALISNT
TESDKYTGASSEKMAAAYAFANGTYIKPMYIKHVFSDSGEFSEFNVGTRAMKET
AYMTEMKTVLSYTGKGNVLAWLPOAGKTGTNTYDEETENIKNTGYVAPDETFF
GYTRKYSMAVWTGYSNRLTPVGGFTVAAKVYR"
BASE COUNT 283 a 214 c 198 g 235 t
ORIGIN
```

Query Match 31.0%; Score 620; DB 1; Length 930;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 770 CTGTGGGATATTACAATACAGCAATACGTTGCCTATCCAGACGATGAATTCGAAGTC 829

|||||

Db 1 CTGTGGGATATTACAATACAGCAATACGTTGCCTATCCAGACGATGAATTCGAAGTC 60

|||||

QY 830 GCTTACCATTTGTGATGTTCTTAACGGTAAAGTCATTCGCCAGCTAGGACGCGCAT 889

|||||

Db 61 GCTTACCATTTGTGATGTTCTTAACGGTAAAGTCATTCGCCAGCTAGGACGCGCAT 120

```
QY 890 CAGTCAAGTAATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCGCGACTGGGA 949
|||||
Db 121 CAGTCAAGTAATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCGCGACTGGGA 180

QY 950 TCAACTATGAACCCGATCACAGACTATGCTCTGCTTGGAGTAGCGTCTACGATTCA 1009
|||||
Db 181 TCAACTATGAACCCGATCACAGACTATGCTCTGCTTGGAGTAGCGTCTACGATTCA 240

QY 1010 ACTGCTACTATGTTTCCAGCATGAGCCCTATAACTACCTCGGACAAATACTCTGTTTAT 1069
|||||
Db 241 ACTGCTACTATGTTTCCAGCATGAGCCCTATAACTACCTCGGACAAATACTCTGTTTAT 300

QY 1070 AACTGGATPAGGGGCTACTTTTGGCAACATCACCTTTCGAATAGCCCTTGAACAATGCGA 1129
|||||
Db 301 AACTGGATPAGGGGCTACTTTTGGCAACATCACCTTTCGAATAGCCCTTGAACAATGCGA 360

QY 1130 AACTGCCAGCCGCTGAAACTCTAAACAAGTCGGAAGTCAACCGCGCGAAGACTTTTCCTA 1189
|||||
Db 361 AACTGCCAGCCGCTGAAACTCTAAACAAGTCGGAAGTCAACCGCGCGAAGACTTTTCCTA 420

QY 1190 AATGGTCTAGGAATCGACTACCAAGTATTCACTACTCAATGCCATTTCAAGTAACACA 1249
|||||
Db 421 AATGGTCTAGGAATCGACTACCAAGTATTCACTACTCAATGCCATTTCAAGTAACACA 480

QY 1250 ACCGAATCAGACAAAAATATATGAGCAAGTAGTAGAAAAGATGGCTGCTTTACGCTGCC 1309
|||||
Db 481 ACCGAATCAGACAAAAATATATGAGCAAGTAGTAGAAAAGATGGCTGCTTTACGCTGCC 540

QY 1310 TTTGCAATGGTGAACCTTACTATAAACAATGTATATCCATAAAGTCGTCTTTAGTGAT 1369
|||||
Db 541 TTTGCAATGGTGAACCTTACTATAAACAATGTATATCCATAAAGTCGTCTTTAGTGAT 600

QY 1370 GGGAGTGAAGAAGAGTTCTCTAATGTCGGAACCTCGTGCATGAAGAAACGACGCCTAT 1429
|||||
Db 601 GGGAGTGAAGAAGAGTTCTCTAATGTCGGAACCTCGTGCATGAAGAAACGACGCCTAT 660

QY 1430 ATGATGACCGA 1440
|||||
Db 661 ATGATGACCGA 671

RESULT 19
AF046237 930 bp DNA linear BCT 30-JUN-1998
LOCUS Streptococcus pneumoniae strain 17619 penicillin-binding protein 1A
DEFINITION (ponA) gene, partial cds.
ACCESSION AF046237
VERSION AF046237.1 GI:2911419
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 930)
AUTHORS Smith,A.M. and Klugman,K.P.
TITLE Alterations in PBP 1A essential for high-level penicillin
resistance in Streptococcus pneumoniae
JOURNAL Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
MEDLINE 98287565
REFERENCE 2 (bases 1 to 930)
AUTHORS Smith,A.M.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg,
Gauteng 2000, South Africa
FEATURES
source
1. .930
Location/Qualifiers
/organism="Streptococcus pneumoniae"
/strain="17619"
/db_xref="taxon:1313"
<1..>930
/gene="ponA"
<1..>930
/gene="ponA"
/feature="PBP 1A"
/codon_start=1
/transl_table=1
/product="penicillin-binding protein 1A"
/protein_id="AAC24700.1"
/db_xref="GI:2911416"
/transl_table="1"
/translation="LMDIYNTDEVYVYPDDELQVASTIVDSNGKVIAQLGARHQSSN
VSGINQAVETNRDWSGTMKPTIDYFALEYGVYDSTATIVHDEPNYPTGNTPVYNN
DRGYGNTLOYLQQRNPVAVETLNKVLNRAKTLFNLGLIDIPDPSIHYSNALISNT
TESDKYTGASSEKMAAAYAFANGTYIKPMYIKHVFSDSGEFSEFNVGTRAMKET
AYMTEMKTVLSYTGKGNVLAWLPOAGKTGTNTYDEETENIKNTGYVAPDETFF
GYTRKYSMAVWTGYSNRLTPVGGFTVAAKVYR"
BASE COUNT 283 a 214 c 198 g 235 t
ORIGIN
```

BASE COUNT 289 a 212 c 196 g 233 t
ORIGIN
/gene="pona"
/note="PBP 1A"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAC24702.1"
/db_xref="GI:2911420"
/translation="LWDIYNDEYVAVPDDELQVASTIVDSNGKVIQALGARHQSNN
VSFGINAVETNRDWSMTKPIDYAPALEYGVYDSTATIVHDEPYNPGNTPVYNN
DRFGVNGITLQALQOQRNPVATLKNVGLNRKATFLNGLIDYPSIHYNAISSNT
TESDKKYAKSEKWAAYAAFAANGTYVKPMYIKHVVSDESEKFEFNVGTRAMKETI
AYMTEHMKTVLSTGTGRNAYLAWLPQAKGTCTSNYDEDEKHIKNGIYVAPDETIV
GYTKRYSNAVMTGYTNRLTPIVIGDFTVAAKVYR"
289 a 212 c 196 g 233 t

Query Match 31.0%; Score 620; DB 1; Length 930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 770 CTGTGGGATATTACAAATACAGCAATACGTTGCTATCCAGACGATGAATTCGAAGTC 829
DB 1 CTGTGGGATATTACAAATACAGCAATACGTTGCTATCCAGACGATGAATTCGAAGTC 60
QY 830 GCTTCTACCATTTGTTGATGTTTCTAACGGTAAAGTCATTGCCAGCTAGGAGCAGCCCAT 889
DB 61 GCTTCTACCATTTGTTGATGTTTCTAACGGTAAAGTCATTGCCAGCTAGGAGCAGCCCAT 120
QY 890 CAGTCAAGTAATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCGCGACGCGGA 949
DB 121 CAGTCAAGTAATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCGCGACGCGGA 180
QY 950 TCAACTATGAACCCGATCACAGCTATGCTCTGCTGGAGTAGCGGTCTACGATTCA 1009
DB 181 TCAACTATGAACCCGATCACAGCTATGCTCTGCTGGAGTAGCGGTCTACGATTCA 240
QY 1010 ACTGCTACTATGCTTACGATGAGCCCTATAACTACCTGGGACAAATACTCCTGTTAT 1069
DB 241 ACTGCTACTATGCTTACGATGAGCCCTATAACTACCTGGGACAAATACTCCTGTTAT 300
QY 1070 AACTGGGATAGGGGCTACTTTGGCAACATCACTTGCATACGCCCTGCAACAATCGGA 1129
DB 301 AACTGGGATAGGGGCTACTTTGGCAACATCACTTGCATACGCCCTGCAACAATCGGA 360
QY 1130 AACGTCGCCAGCGGTGAAACTCTAACAGAGTCGGACTCAACCGCGCCAGACTTTCCCTA 1189
DB 361 AACGTCGCCAGCGGTGAAACTCTAACAGAGTCGGACTCAACCGCGCCAGACTTTCCCTA 420
QY 1190 AATGGTCTAGGAATCGACTACCAAGTATTAATCAATGCAATTTCAAGTAACACA 1249
DB 421 AATGGTCTAGGAATCGACTACCAAGTATTAATCAATGCAATTTCAAGTAACACA 480
QY 1250 ACCGAATCAGACAAAATATATGGAGCAAGTAGTGAAGAGTGGCTGCTTACGCTGCC 1309
DB 481 ACCGAATCAGACAAAATATATGGAGCAAGTAGTGAAGAGTGGCTGCTTACGCTGCC 540
QY 1310 TTGTCAAATGGTGAACCTTACTATTAACCAATGTATTCATTAAGTCTCTTTAGTGAT 1369
DB 541 TTGTCAAATGGTGAACCTTACTATTAACCAATGTATTCATTAAGTCTCTTTAGTGAT 600
QY 1370 GGGAGTGAAGAAAGAGTTCTTAATGTCGGAACCTGCTGCATTAAGGAAACGACGCTAT 1429
DB 601 GGGAGTGAAGAAAGAGTTCTTAATGTCGGAACCTGCTGCATTAAGGAAACGACGCTAT 660
QY 1430 ATGATGACCGA 1440
DB 661 ATGATGACCGA 671

RESULT 20
AX111320
LOCUS 930 bp DNA linear PAT 30-APR-2001

DEFINITION Sequence 2053 from Patent WO0123604.
ACCESSION AX111320
VERSION AX111320.1 GI:13927612
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
REFERENCE 1 (bases 1 to 930)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2053 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
Location/Qualifiers
1..930
/organism="Streptococcus pneumoniae"
/strain="17619"
/db_xref="taxon:1313"
BASE COUNT 289 a 212 c 196 g 233 t
ORIGIN

Query Match 31.0%; Score 620; DB 6; Length 930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 770 CTGTGGGATATTACAAATACAGCAATACGTTGCTATCCAGACGATGAATTCGAAGTC 829
DB 1 CTGTGGGATATTACAAATACAGCAATACGTTGCTATCCAGACGATGAATTCGAAGTC 60
QY 830 GCTTCTACCATTTGTTGATGTTTCTAACGGTAAAGTCATTGCCAGCTAGGAGCAGCCCAT 889
DB 61 GCTTCTACCATTTGTTGATGTTTCTAACGGTAAAGTCATTGCCAGCTAGGAGCAGCCCAT 120
QY 890 CAGTCAAGTAATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCGCGACGCGGA 949
DB 121 CAGTCAAGTAATGTTTCTTCGGAATTAACCAAGCAGTAGAACAACCGCGACGCGGA 180
QY 950 TCAACTATGAACCCGATCACAGCTATGCTCTGCTGGAGTAGCGGTCTACGATTCA 1009
DB 181 TCAACTATGAACCCGATCACAGCTATGCTCTGCTGGAGTAGCGGTCTACGATTCA 240
QY 1010 ACTGCTACTATGCTTACGATGAGCCCTATAACTACCTGGGACAAATACTCCTGTTAT 1069
DB 241 ACTGCTACTATGCTTACGATGAGCCCTATAACTACCTGGGACAAATACTCCTGTTAT 300
QY 1070 AACTGGGATAGGGGCTACTTTGGCAACATCACTTGCATACGCCCTGCAACAATCGGA 1129
DB 301 AACTGGGATAGGGGCTACTTTGGCAACATCACTTGCATACGCCCTGCAACAATCGGA 360
QY 1130 AACGTCGCCAGCGGTGAAACTCTAACAGAGTCGGACTCAACCGCGCCAGACTTTCCCTA 1189
DB 361 AACGTCGCCAGCGGTGAAACTCTAACAGAGTCGGACTCAACCGCGCCAGACTTTCCCTA 420
QY 1190 AATGGTCTAGGAATCGACTACCAAGTATTAATCAATGCAATTTCAAGTAACACA 1249
DB 421 AATGGTCTAGGAATCGACTACCAAGTATTAATCAATGCAATTTCAAGTAACACA 480
QY 1250 ACCGAATCAGACAAAATATATGGAGCAAGTAGTGAAGAGTGGCTGCTTACGCTGCC 1309
DB 481 ACCGAATCAGACAAAATATATGGAGCAAGTAGTGAAGAGTGGCTGCTTACGCTGCC 540
QY 1310 TTGTCAAATGGTGAACCTTACTATTAACCAATGTATTCATTAAGTCTCTTTAGTGAT 1369
DB 541 TTGTCAAATGGTGAACCTTACTATTAACCAATGTATTCATTAAGTCTCTTTAGTGAT 600
QY 1370 GGGAGTGAAGAAAGAGTTCTTAATGTCGGAACCTGCTGCATTAAGGAAACGACGCTAT 1429
DB 601 GGGAGTGAAGAAAGAGTTCTTAATGTCGGAACCTGCTGCATTAAGGAAACGACGCTAT 660
QY 1430 ATGATGACCGA 1440

```
Db 661 ATGATGACCGA 671
|||||
RESULT 21
AX110281
LOCUS AX110281 1199 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1014 from Patent WO0123604.
ACCESSION AX110281
VERSION AX110281.1 GI:13926573
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
REFERENCE 1 (bases 1 to 1199)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 1014 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
source
Location/Qualifiers
1..1199
/organism="Streptococcus pneumoniae"
/strain="StrR-11"
/db_xref="taxon:1313"
BASE COUNT 386 a 286 c 240 g 287 t
ORIGIN
Query Match 29.5%; Score 590; DB 6; Length 1199;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 690; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 761 CAAAACATCTGTGGATATTACATACAGACGATAGTTGCCCTATCCAGACGATGAA 820
Db 1 CAAAACATCTGTGGATATTACAAATACAGACGATAGTTGCCCTATCCAGACGATGAA 60
QY 821 TTGCAAGTCGCTTCTACCATGTTGATGTTCTTAACGGTAAAGTCAATGTCGCCAGTAGGA 880
Db 61 TTGCAAGTCGCTTCTACCATGTTGATGTTCTTAACGGTAAAGTCAATGTCGCCAGTAGGA 120
QY 881 GCACGCCATCAGTCAAGTAAATGTTCCCTTCGGAAATTAACCAAGCAGTAGAACAACCCG 940
Db 121 GCACGCCATCAGTCAAGTAAATGTTCCCTTCGGAAATTAACCAAGCAGTAGAACAACCCG 180
QY 941 GACTGGGGATCAACTATGAACCGATACAGACTATGCTCCTGCTTGGAGTACGGTGC 1000
Db 181 GACTGGGGATCAACTATGAACCGATACAGACTATGCTCCTGCTTGGAGTACGGTGC 240
QY 1001 TAGGATTCAACTGCTACTATCGTTACGATGAGCCCTATAAATACCTTGGGACAAATACT 1060
Db 241 TAGGATTCAACTGCTACTATCGTTACGATGAGCCCTATAAATACCTTGGGACAAATACC 300
QY 1061 CCGTTTATACTGGGATAGGGGCTACTTTGGCAACATCACCTTGGCAATACGCCCTGCAA 1120
Db 301 CCGTTTATACTGGGATAGGGGCTACTTTGGCAACATCACCTTGGCAATACGCCCTGCAA 360
QY 1121 CAATCCGAAACGTCGCCCGGAACTCTAAACAAGGTCGGACTCAACCCGCCCAAG 1180
Db 361 CAATCCGAAACGTCGCCCGGAACTCTAAACAAGGTCGGACTCAACCCGCCCAAG 420
QY 1181 ACTTTCTAAATGGTCTAGGAATCGACTACCCAAATATTACACTACATGCAATGCAATTTCA 1240
Db 421 ACTTTCTAAATGGTCTAGGAATCGACTACCCAAATATTACACTACATGCAATGCAATTTCA 480
QY 1241 AGTAACACACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGCTGCT 1300
Db 481 AGTAACACACCGAATCAGACAAAAAATATGGAGCAAGTAGTGAAGAAGATGGCTGCTGCT 540
QY 1301 TAGGCTGCCTTGGCAATGGTGAACCTTACTATAAACCAATGATATATCAATAAAGTCGTC 1360
|||||
```

```
Db 541 TACGCTGCCCTTGCATAATGCTGGAACCTACTATAAACCAATGATATATCCATAAAGTCGTC 600
QY 1361 TTTAGTGATGGAGTGAAAAAGAGTTCTCTAAATGTCGGAACCTGCTGCATCAAGGAAACG 1420
|||||
Db 601 TTTAGTGATGGAGTGAAAAAGAGTTCTCTAAATGTCGGAACCTGCTGCATCAAGGAAACG 660
QY 1421 ACAGCCTATATGATGACCCGACATGATGAAAC 1452
Db 661 ACAGCCTATATGATGACCCGACATGATGAAAC 692
|||||
RESULT 22
AX110284
LOCUS AX110284 1229 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1017 from Patent WO0123604.
ACCESSION AX110284
VERSION AX110284.1 GI:13926576
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
REFERENCE 1 (bases 1 to 1229)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 1017 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
source
Location/Qualifiers
1..1229
/organism="Streptococcus pneumoniae"
/strain="StrR-14"
/db_xref="taxon:1313"
BASE COUNT 399 a 287 c 240 g 303 t
ORIGIN
Query Match 28.4%; Score 567; DB 6; Length 1229;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 667; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 745 TTGTAGCAAGAAGCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATACGTTGC 804
Db 1 TTGTAGCAAGAAGCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATACGTTGC 60
QY 805 CTATCAGACGATGAATTGCAAGTCGCTTCTACCATGTTGATGTTCTTAACGTAAGT 864
Db 61 CTATCAGACGATGAATTGCAAGTCGCTTCTACCATGTTGATGTTCTTAACGTAAGT 120
QY 865 CATTTGCCAGCTAGGAGCAGCCCATCAGTCAAGTAATGTTTCCCTTCGGAATTAACCAAGC 924
Db 121 CATTTGCCAGCTAGGAGCAGCCCATCAGTCAAGTAATGTTTCCCTTCGGAATTAACCAAGC 180
QY 925 AGTAGAACAACCCGCGACTGGGGATCAACTATGAACCCGATCAGACACTATGCTCTCTGC 984
Db 181 AGTAGAACAACCCGCGACTGGGGATCAACTATGAACCCGATCAGACACTATGCTCTCTGC 240
QY 985 CTTGGAGTACGGTGTCTACGATTCAACTGCTACTACTGTTTCAGATGAGCCCTATAACTA 1044
Db 241 CTTGGAGTACGGTGTCTACGATTCAACTGCTACTACTGTTTCAGATGAGCCCTATAACTA 300
QY 1045 CCCTGGGACAAATACTCCTGTTTATACTGGGATAGGGGCTACTTTTGGCAACATCACCTT 1104
Db 301 CCCTGGGACAAATACTCCTGTTTATACTGGGATAGGGGCTACTTTTGGCAACATCACCTT 360
QY 1105 GCAATACGCCCTGCACAATCGGAAACGTCGCCAGCGGTGGAACTCTAAACAAGGTGCG 1164
Db 361 GCAATACGCCCTGCACAATCGGAAACGTCGCCAGCGGTGGAACTCTAAACAAGGTGCG 420
QY 1165 ACTCAACCCGCCCAAGACTTCTCTAAATGCTTAGGAATCGACTACCCAGTATTACCTA 1224
Db 421 ACTCAACCCGCCCAAGACTTCTCTAAATGCTTAGGAATCGACTACCCAGTATTACCTA 480
|||||
```

```
QY 1225 CTCAATGCCATTTCAGTAACACAAACCGAATCAGACAAAAAATATGGAGCAAGTAGTGA 1284
|||||
Db 481 CTCAATGCCATTTCAGTAACACAAACCGAATCAGACAAAAAATATGGAGCAAGTAGTGA 540
|||||
QY 1285 AAGATGGCTGCTGCTTACGTCGCCCTTTGCAATGTGGAACCTACTATTAACCAATGTA 1344
|||||
Db 541 AAGATGGCTGCTGCTTACGTCGCCCTTTGCAATGTGGAACCTACTATTAACCAATGTA 600
|||||
QY 1345 TATCCATAAAGTCGCTTTAGTAGTGGAGTGGAGTAAAAAGAGTTCTCTAATGTGCGAACTCG 1404
|||||
Db 601 TATCCATAAAGTCGCTTTAGTAGTGGAGTGGAGTAAAAAGAGTTCTCTAATGTGCGAACTCG 660
|||||
QY 1405 TCCCATGAA 1413
|||||
Db 661 TGCCATGAA 669
|||||

RESULT 23
AF046236
LOCUS
DEFINITION Streptococcus pneumoniae strain N94 penicillin-binding protein 1A
ACCESSION AF046236
VERSION AF046236.1 GI:2911417
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 930)
Smith,A.M. and Klugman,K.P.
Alterations in pBP 1A essential for high-level penicillin
resistance in Streptococcus pneumoniae
JOURNAL Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
MEDLINE 98287565
REFERENCE Smith,A.M.
Direct Submission
TITLE Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg,
JOURNAL Gauteng 2000, South Africa
FEATURES
source
Location/Qualifiers
1..930
/organism="Streptococcus pneumoniae"
/strain="N94"
/db_xref="taxon:1313"
<1..>930
/gene="ponA"
/feature="ponA"
/feature="ponA"
/feature="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAC24701.1"
/db_xref="GI:2911418"
/translation="LWDIYNTDEVYVYPDELOVASTIVDSNGKVIQALGARHOSN
VSFGINQAVETNRDWSMTKPTIDVAPALEYGVXDSTATIVHDEPNYPGNTVPYNN
DRGYFNTILQVALQSRNPVAVETLNKVLNRAKTFNLGLIDYPSIHYSNAISSNT
TESDKYGSASEKMAAAYAAAFANGGTIKPMYTHKVVFSDSGERFSVNGTAMKETT
AYMTMKTKVLVYIGRGAYLPLPQAGKTGTSYTDDEIEKYIKNTGYVAPDEMVF
GYTRKYAMAVMTGYNRLTPVVGDTVAAKYR"
BASE COUNT 287 a 213 c 201 g 229 t
ORIGIN

Query Match 27.6%; Score 551; DB 1; Length 930;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 651; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 770 CTCTGGGATATTACAAATACAGACGAATACGTTGCCATCCAGACCATGAATTCGAAGTC 829
|||||
Db 1 CTCTGGGATATTACAAATACAGACGAATACGTTGCCATCCAGACCATGAATTCGAAGTC 60
|||||
```

```
QY 830 GCTTCTACCATTGTTGATGTTTCTAACGGTAAAGTCAATGCCAGCTAGGAGCACGCCAT 889
|||||
Db 61 GCTTCTACCATTGTTGATGTTTCTAACGGTAAAGTCAATGCCAGCTAGGAGCACGCCAT 120
|||||
QY 890 CAGTCAAGTAATGTTTCCCTTCGGAAATTAACCAAGCAGTAGAACAACACCGAGCTGGGA 949
|||||
Db 121 CAGTCAAGTAATGTTTCCCTTCGGAAATTAACCAAGCAGTAGAACAACACCGAGCTGGGA 180
|||||
QY 950 TCAACTATGAACCGCATCACAGACTATGCTCCCTGCGCTTGGAGTACGCTGCTACGATTCA 1009
|||||
Db 181 TCAACTATGAACCGCATCACAGACTATGCTCCCTGCGCTTGGAGTACGCTGCTACGATTCA 240
|||||
QY 1010 ACTGCTACTATCGTTCACGATGAGCCCTATACTACCTGGGACAAATACTCTCTGTTAT 1069
|||||
Db 241 ACTGCTACTATCGTTCACGATGAGCCCTATACTACCTGGGACAAATACTCTCTGTTAT 300
|||||
QY 1070 AACTGGGATAGGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACATCGGA 1129
|||||
Db 301 AACTGGGATAGGGCTACTTTGGCAACATCACCTTGCATACGCCCTGCAACATCGGA 360
|||||
QY 1130 AAGCTCCAGCCGCTGAAACTCTAAACAGGTGCGACTCAACCGCGCAAGACTTTCCCTA 1189
|||||
Db 361 AAGCTCCAGCCGCTGAAACTCTAAACAGGTGCGACTCAACCGCGCAAGACTTTCCCTA 420
|||||
QY 1190 AATGCTCTAGGAATCGACTACCAAGTATTCACTACTCAAAATGCCATTTCAAGTAACACA 1249
|||||
Db 421 AATGCTCTAGGAATCGACTACCAAGTATTCACTACTCAAAATGCCATTTCAAGTAACACA 480
|||||
QY 1250 ACGGAATCAGACAAAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTGCTAGCTGCC 1309
|||||
Db 481 ACGGAATCAGACAAAAAATATGGAGCAAGTAGTGAAGAAGTGGCTGCTGCTAGCTGCC 540
|||||
QY 1310 TTTCGAAATGGTGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGAT 1369
|||||
Db 541 TTTCGAAATGGTGGAACTTACTATAAACCAATGTATATCCATAAAGTCGCTTTAGTGAT 600
|||||
QY 1370 GGGAGTGAAGAAGAGTTCTCTAATGTCGGAACCTCGTCCCATGAAGGAACGAC 1422
|||||
Db 601 GGGAGTGAAGAAGAGTTCTCTAATGTCGGAACCTCGTCCCATGAAGGAACGAC 653
|||||

RESULT 24
AF046234
LOCUS
DEFINITION Streptococcus pneumoniae strain 56739 penicillin-binding protein 1A
ACCESSION AF046234
VERSION AF046234.1 GI:2911413
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
1 (bases 1 to 930)
Smith,A.M. and Klugman,K.P.
Alterations in pBP 1A essential for high-level penicillin
resistance in Streptococcus pneumoniae
JOURNAL Antimicrob. Agents Chemother. 42 (6), 1329-1333 (1998)
MEDLINE 98287565
REFERENCE Smith,A.M.
Direct Submission
TITLE Submitted (06-FEB-1998) Microbiology, SAIMR Hospital, Johannesburg,
JOURNAL Gauteng 2000, South Africa
FEATURES
source
Location/Qualifiers
1..930
/organism="Streptococcus pneumoniae"
/strain="56739"
/db_xref="taxon:1313"
<1..>930
/feature="ponA"
/feature="ponA"
/feature="ponA"
/feature="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAC24701.1"
/db_xref="GI:2911418"
/translation="LWDIYNTDEVYVYPDELOVASTIVDSNGKVIQALGARHOSN
VSFGINQAVETNRDWSMTKPTIDVAPALEYGVXDSTATIVHDEPNYPGNTVPYNN
DRGYFNTILQVALQSRNPVAVETLNKVLNRAKTFNLGLIDYPSIHYSNAISSNT
TESDKYGSASEKMAAAYAAAFANGGTIKPMYTHKVVFSDSGERFSVNGTAMKETT
AYMTMKTKVLVYIGRGAYLPLPQAGKTGTSYTDDEIEKYIKNTGYVAPDEMVF
GYTRKYAMAVMTGYNRLTPVVGDTVAAKYR"
BASE COUNT 287 a 213 c 201 g 229 t
ORIGIN
```



```

/gene="ponA"
/note="PBP 1A"
/codeon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAC24599.1"
/db_xref="GI:2911414"
/translation="LWDIYNTDEYVAYPDELOVASTIVDSNGKVIQAQIGARHQSSN
VSFGINQAVETNRDWSGTWKPPIPDYAPALEYGVYDSTATIVHDPEYVNPCTNPVNW
DRGYINFTLOYALQDSRRNPVAYETLNKLRQAKPTFLNGLIGDYPHSIHVSN1ISNT
TESDDKYGASSEKMAAYAAAFANGTYTYKPMYIHKVVFSGSEKESFDAGTRAMKET
AYMTKMTKTVLYGIRGAYFLPWLQAGKTGTSNTDDEIEK1IANTGIVYAPDEMFFV
GYMKYAMAYVTGYSNRLTPLVGVNGLUTVAAKVYR"
288 a 216 c 196 g 230 t

```

Query Match	23.1%	Score 461	DB 1	Length 930
Best Local Similarity	99.5%	Pred. No. 2.3e-248		
Matches 611	Conservative 0	Mismatches 3	Indels 0	Gaps 0
QY	770	CTGTGGGATATTTACAATACAGACGAATACGTTGCCTATCCAGACGATGAATTCGAAGTC	829	
DB	1	CTGTGGGATATTTACAATACAGACGAATACGTTGCCTATCCAGACGATGAATTCGAAGTC	60	
QY	830	CGTTCTACCATGTTGTGATGTTTCTAACGGTAAAGTCAATGCCCCAGCTAGGAGCACGGCAT	889	
DB	61	CGTTCTACCATGTTGTGATGTTTCTAACGGTAAAGTCAATGCCCCAGCTAGGAGCACGGCAT	120	
QY	890	CAGTCAAGTAAATGTTCTTCGGAAATTTAACCAAGCAGTAAAGAACAAACCGGACTGGGGA	949	
DB	121	CAGTCAAGTAAATGTTCTTCGGAAATTTAACCAAGCAGTAAAGAACAAACCGGACTGGGGA	180	
QY	950	TCAACTATGAACCGGATCAGACGATATGCTCTGCCCTTGGAGTAGCGGTGCTACGATTCA	1009	
DB	181	TCAACTATGAACCGGATCAGACGATATGCTCTGCCCTTGGAGTAGCGGTGCTACGATTCA	240	
QY	1010	ACTGCTACTATCGTTTCACGATGAGCCCTATAAATACCTACCTGGGACAAATACCTCTGTTTAT	1069	
DB	241	ACTGCTACTATCGTTTCACGATGAGCCCTATAAATACCTACCTGGGACAAATACCTCTGTTTAT	300	
QY	1070	AACGTGGGATAGGGCTACTTTTGGCAACATCACCTTGGCAATACGCGCTTGCAACAAATCGCGA	1129	
DB	301	AACGTGGGATAGGGCTACTTTTGGCAACATCACCTTGGCAATACGCGCTTGCAACAAATCGCGA	360	
QY	1130	AAGTCTCCACGCGTGGAAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTTCCTA	1189	
DB	361	AAGTCTCCACGCGTGGAAACTCTAAACAAGTCGGACTCAACCGCGCAAGACTTTTCCTA	420	
QY	1190	AATGGTCTAGGAATCGACTACCCAAAGTATTCAGTACTCAAAATGCGCATTTCAAGTAAACACA	1249	
DB	421	AATGGTCTAGGAATCGACTACCCAAAGTATTCAGTACTCAAAATGCGCATTTCAAGTAAACACA	480	
QY	1250	ACCGAATCAGACAAAAAATATGAGCAAGTAGTGAAAGATGCGTGTCTTACGCTGCC	1309	
DB	481	ACCGAATCAGACAAAAAATATGAGCAAGTAGTGAAAGATGCGTGTCTTACGCTGCC	540	
QY	1310	TTTGCAAATGGTGAACCTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTAGTGAT	1369	
DB	541	TTTGCAAATGGTGAACCTTACTATAAACCAATGTATATCCATAAAGTCGCTTTTAGTGAT	600	
QY	1370	GGGAGTAAAAAAGA 1383		
DB	601	GGGAGTAAAAAAGA 614		
RESULT	25			
AF210747				
LOCUS	AF210747	2160 bp	DNA	linear
DEFINITION	Streptococcus pneumoniae isolate SP 1465 penicillin-binding protein			
ACCESSION	AF210747			
VERSION	AF210747.1	GI:6563340		

RESULT	25
AF210747	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	

KEYWORDS	Streptococcus pneumoniae.
SOURCE	Streptococcus pneumoniae
ORGANISM	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.
REFERENCE	1 (bases 1 to 2160)
AUTHORS	Ferroni, A. and Berche, P.
TITLE	Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children
JOURNAL	J. Med. Microbiol. 50 (9), 828-832 (2001)
MEDLINE	21432820
PUBMED	11549185
REFERENCE	2 (bases 1 to 2160)
AUTHORS	Ferroni, A. and Berche, P.
TITLE	Direct Submission
JOURNAL	Submitted (03-DEC-1999) Microbiology, Hospital Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France
FEATURES	Location/Qualifiers
source	1. .2160 /organism="Streptococcus pneumoniae" /serotype="23F" /isolate="SP 1465" /db_xref="taxon:1313" 1. .2160 /gene="pbpla" 1. .2160 /gene="pbpla" /codon_start=1 /transl_table=11 /product="penicillin-binding protein 1A" /protein_id="AAFI1727.1" /db_xref="GI:6563341" /translation="MKKPTLRILKYLISFSLVIAAIVLGGVFYVYSKAPLSLE KIVATTSKIIYDNKNLSIADLGSERRVNAQANDPTDLVKAIVSIDBHRFFDHRGID TIRILGAFNRNOSLOGGSTITQOLI KLT YFSTSDOT ISRKAQEWLAIQLEQ ATKEILTYIVNKVYNSNGYMGOTAAQNYGKDLRELSLPOLALLAGMPAQNDYD YSHQEDLRNRLVLSMKQGYISAPQYKAKINTPTDGLQSLKSVNSYPAYMAYLV KEVIDQEQETGNLLTGMDVTINVDQEAQKHLWDIYNSDQIVSYDPDDQLQVASTVV DVNSGVIAQLGARHQAASNVSFGTNQAVETNRDQSNVTKMPTIDYAPAEIYGVYDSTAT MVNDIYNYETGTPSTNYDRAVYFGNITLQALQQSRNPVAVETLNKVGDLDRAKFN GLGIDYDMHYSNAISNTTYESNKQYCASSEKMAAFAAFANGIYHKPMYINKIVFS DGSEKFSFDAGTRMKETATTYEMKMTKTVLSYGTGRNAVLAWLPOAQKGTGTSNYDT EIEENHIKTSQFVAPDELFCAGYTRKYSMAVVTGYSNRLTPLYVGNGLTVAAKVYRSMYI LSSGSNPEDNIEPGLYRNGEYFVKNGARSTWSSPAQPQPPSTSSSSSDSSTSSQSS STPTSTNNSTTTTNNNTQSNTPPDQONQPAQP"
BASE COUNT	692 a 505 c 405 g 558 t
ORIGIN	

Query Match	21.8%	Score 436;	DB 1;	Length 2160;
Best Local Similarity	99.6%;	Pred. No. 3e-234;		
Matches 536;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY 1462	TTATGGAACTGGAGCAAAATGCCTATCTTGGCTCCCTCAGGCTGGTAAACAGGAAC	1521		
Db 1620	TTATGGAACTGGAGCAAAATGCCTATCTTGGCTCCCTCAGGCTGGTAAACAGGAAC	1679		
QY 1522	CTCTAACTATACAGACGAGGAAATTTGAAACCCACATCAAGACCTCTCAATTTGTAGCACC	1581		
Db 1680	CTCTAACTATACAGACGAGGAAATTTGAAACCCACATCAAGACCTCTCAATTTGTAGCACC	1739		
QY 1582	TGATGAACATATTGCTGGCTATACGCGTAAATATTCAATGGCTGTATGGACAGGCTATTTC	1641		
Db 1740	TGATGAACATATTGCTGGCTATACGCGTAAATATTCAATGGCTGTATGGACAGGCTATTTC	1799		
QY 1642	TAAACCGTCTGACACCACTTTGTAGCAATATGGCTTACGGTCTGCTGCCAAAGTTTACC GGTC	1701		
Db 1800	TAAACCGTCTGACACCACTTTGTAGCAATATGGCTTACGGTCTGCTGCCAAAGTTTACC GGTC	1859		
QY 1702	TATGATGACCTACTGTCTGAAGGAAGCAATCCAGAGATTTGGAATATACAGAGGGGCT	1761		
Db 1860	TATGATGACCTACTGTCTGAAGGAAGCAATCCAGAGATTTGGAATATACAGAGGGGCT	1919		

Qy	1762	CTACAGAAATGGAGAAATTCGTATTATTAATAATGGTGCCTGTTCTACGTGGAAACTCACCTGC	1821
Db	1920	CTACAGAAATGGAGAAATTCGTATTATTAATAATGGTGCCTGTTCTACGTGGAGCTCACCTGC	1979
Qy	1822	TCCACAACACCCCCCACTCAACTGAAAGTTCGAAGTCTCAAGTCTCATCATGATATGTTCAACTTCACA	1881
Db	1980	TCCACAACACCCCCCACTCAACTGAAAGTTCGAAGTCTCAAGTCTCATCATGATATGTTCAACTTCACA	2039
Qy	1882	GTCTAGTCTCAACCACTCCCAAGCACAAATAATAGTACGACTACCAATCCTCAACAATAATAC	1941
Db	2040	GTCTAGTCTCAACCACTCCCAAGCACAAATAATAGTACGACTACCAATCCTCAACAATAATAC	2099
Qy	1942	GCACAATCAAAATCAACCCCTGATCAACAATAATCAGAAATCCTCAACGAGCACAAACCA	1999
Db	2100	GCAACAATCAAAATCAACCCCTGATCAACAATAATCAGAAATCCTCAACGAGCACAAACCA	2157
RESULT	26		
SPNEU1914/c			
LOCUS	SPNEU1914	30457 bp	DNA linear
DEFINITION	Streptococcus pneumoniae clone G54, *** SEQUENCING IN PROGRESS ***, in ordered pieces.		
ACCESSION	AL449936		
VERSION	AL449936.1	GI:11545161	
KEYWORDS	HTG; HTGS_PHASE2.		
SOURCE	Streptococcus pneumoniae.		
ORGANISM	Streptococcus pneumoniae		
	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
REFERENCE	Streptococcus.		
AUTHORS	1 (bases 1 to 30457) Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Humbert,Y., Friedli,L., Guerrier,M., Grand-Schenk,E., Gandin,C., de Francesco,M., Polissi,A., Buell,G., Feger,G., Garcia,E., Peitsch,M., and Garcia-Bustos,J.F.		
TITLE	Annotated draft genomic sequence from a Streptococcus pneumoniae type 19F clinical isolate		
JOURNAL	Microb. Drug Resist.	7 (2), 99-125	(2001)
MEDLINE	2133529		
REFERENCE	2 (bases 1 to 30457)		
AUTHORS	Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Polissi,A., Humbert,Y., Friedli,L., Guerrier,M., Grand-Schenk,E., Gandin,C., de Francesco,M., Buell,G., Feger,G., Garcia,E., Peitsch,M. and Garcia-Bustos,J.F.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-Oct-2000)	Research Department, Glaxo Wellcome, S.A., Severo Ochoa 2, 28760 Tres Cantos, SPAIN	
COMMENT	* NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved.		
FEATURES	Location/Qualifiers		
source	1. 30457		
	/organism="Streptococcus pneumoniae" /serotype="19F" /db_xref="taxon:1313" /clone="G54"		
BASE COUNT	9377 a	5459 c	6906 g
ORIGIN	8705 t	10 others	
Query Match	21.8%	Score 436:	DB 2: Length 30457;
Best Local Similarity	99.6%	Pred. No. 2.9e-234;	
Matches 536:	Conservative	0: Mismatches	2: Indels 0: Gaps 0:
Qy	1462	TTATGGAACTGGACGAATGCCTATCTTGCTTGCTCCCTCAGGCTGGTAAACAGGAAC	1521
Db	8148	TTATGGAACTGGACGAATGCCTATCTTGCTTGCTCCCTCAGGCTGGTAAACAGGAAC	8089
Qy	1522	CTCTAATATACACGAGGAAATGAAACCAATCAAGACCTCTCAATTTGTAGACC	1581
Db	8088	CTCTAATATACACGAGGAAATGAAACCAATCAAGACCTCTCAATTTGTAGACC	8029

Qy	1582	TGATGAACCTATTGCTGGCTATACGCGTAATAATTCAATGGCTGTATGGACAGGCTATTTC	1641
Db	8028	TGATGAACCTATTGCTGGCTATACGCGTAATAATTCAATGGCTGTATGGACAGGCTATTTC	7969
Qy	1642	TAACCGTCTGACACCACTTGTAGGCAATGGCCTTTACGGTCGCTGCCAAAGCTTTACCGCTC	1701
Db	7968	TAACCGTCTGACACCACTTGTAGGCAATGGCCTTTACGGTCGCTGCCAAAGCTTTACCGCTC	7909
Qy	1702	TATGATGACCTTACCTGTCTGAAGGAAGCAATCCAGAGAAGATTGGAATATACGAGAGGGCT	1761
Db	7908	TATGATGACCTTACCTGTCTGAAGGAAGCAATCCAGAGAAGATTGGAATATACGAGAGGGCT	7849
Qy	1762	CTACAGAAATGGAGAATTCGATTTTAAAAATGGTGCCTTCTACGTGGAACCTCACCTGC	1821
Db	7848	CTACAGAAATGGAGAATTCGATTTTAAAAATGGTGCCTTCTACGTGGAACCTCACCTGC	7789
Qy	1822	TCCACAACACCCCATCAACTGAAAGTTCAAGCTCATCATCAGATAGTTCAACTTCACA	1881
Db	7788	TCCACAACACCCCATCAACTGAAAGTTCAAGCTCATCATCAGATAGTTCAACTTCACA	7729
Qy	1982	GTCTAGTCTAACCACTCCACGACACAATAATAGTAGGACTTACCAATCCTTAACAATAATAC	1941
Db	7728	GTCTAGTCTAACCACTCCACGACACAATAATAGTAGGACTTACCAATCCTTAACAATAATAC	7669
Qy	1942	GCAACATCAAAATACAACCCCTGATCAACAAAATCAGAATCCTCAACGACACACAACA	1999
Db	7668	GCAACATCAAAATACAACCCCTGATCAACAAAATCAGAATCCTCAACGACACACAACA	7611
RESULT	27		
AF210752			
LOCUS	AF210752	2160 bp DNA linear BCT 17-SEP-2001	
DEFINITION		Streptococcus pneumoniae isolate BM 4200 penicillin-binding protein 1A (pbp1a) gene, complete cds.	
ACCESSION	AF210752		
VERSION	AF210752.1	GI:6563350	
KEYWORDS			
SOURCE		Streptococcus pneumoniae.	
ORGANISM		Streptococcus pneumoniae.	
REFERENCE		Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;	
AUTHORS		Streptococcus.	
TITLE		1 (bases 1 to 2160)	
JOURNAL		Ferroni,A. and Berche,P.	
MEDLINE		Alterations to penicillin-binding proteins 1A, 2B and 2X amongst	
PUBMED		penicillin-resistant clinical isolates of Streptococcus pneumoniae	
REFERENCE		serotype 23F from the nasopharyngeal flora of children	
AUTHORS		J. Med. Microbiol. 50 (9), 828-832 (2001)	
TITLE		21432820	
JOURNAL		11549185	
MEDLINE		2 (bases 1 to 2160)	
PUBMED		Ferroni,A. and Berche,P.	
REFERENCE		Direct Submission	
AUTHORS		Submitted (03-DEC-1999) Microbiology, Hospital	
TITLE		Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France	
JOURNAL		Location/Qualifiers	
FEATURES		1. .2160	
source		/organism="Streptococcus pneumoniae"	
		/serotype="23F"	
		/isolate="BM 4200"	
		/db_xref="taxon:1313"	
		/note="strain isolated in 1978 in Paris"	
gene		1. .2160	
		/gene="pbp1a"	
CDS		1. .2160	
		/gene="pbp1a"	
		/codon_start=1	
		/transl_table=11	
		/product="penicillin-binding protein 1A"	
		/protein_id="AAE17262.1"	
		/db_xref="GI:6563351"	
		/translation="MNKPTIRLILKYLSISFLSLVIAAIVLGGVFFVYVSKAPSLSE	
		SKIVATFSKIDNKQLADLGRSVAQANDIPTDLVKAIVSDIHRDFHRGID	
		TRILCAAFKRNLOSNGSLSTTOLIKLTVSTSDTISRKAQEAWLAIQLEOK	

ATKQBELIYYINKVMSNGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPOAPNOYDP
YSHPEAQRDRNLRLYSEMKNOGYISAEQYKAVNPTITDGLQSLKSASNYPAYMDNLY
KEVINQVEEETGYNLLTMDVYTNVDOEAQKHLWDIYNTDEYVAYPDEQVASTIV
DVSNGKVIQAOLCARHOSNSVSGFINOAVENTNRDWSGTMKPTIDYAPALEGYVYDSTAT
IVHDEPYNYPGTDPVYNDRGYSFGNITLOVALOQSRNVPATVETLNKVCINLAKRTFLN
GLGIDYPSLHYSNALSSNTTSDSKYKYGASSEKMAAAYAFANGYTYYPWYHKKYVFS
DGSEKFSNVTGRANKETATMTDMKTVILVYIGIRGAYILPWLPOAKRTGTSNVTDE
LSEKIKNTGYVADPEMEVGYTRKYAMAVWTGYSNRLPLVGDGLUTVAAKYVRSMTY
LSEGSNPDWNIPEGLYRNGFEVKNGARSTWSSPAQPOPPSTESSSSSSSSSTSQSN
STTPTSTNNSTTTNNNTTQQSNNTTPOQNONPQAPQ*
BASE COUNT 698 a 520 c 416 g 526 t
ORIGIN

Query Match 21.6%; Score 432; DB 1; Length 2160;
Best Local Similarity 98.9%; Pred. No. 5.4e-232;
Matches 1032; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 TAAATCTACGACATAAATAAATCAACTCATTCGTGACTTGGGTTCTGAACGGCGCGTCAA 60
DB 159 TAAATCTACGACATAAATAAATCAACTCATTCGTGACTTGGGTTCTGAACGGCGCGTCAA 218
QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
DB 219 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACCACAGGGGATGATACCATCCGTATCCTGGGAGCTTCTTCGCGCAA 180
DB 279 TCGCTTCTTCGACCACAGGGGATGATACCATCCGTATCCTGGGAGCTTCTTCGCGCAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTFCACCCACAGTGTGATTAAGTTGAC 240
DB 339 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTFCACCCACAGTGTGATTAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTCTAGGCTCAGSAGCTTGGTT 300
DB 399 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTCTAGGCTCAGSAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
DB 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
QY 361 GGTCTACATGCTAATGGGAACATATGGAATGCAGACAGCAGCTCAAACTACTATGGTAA 420
DB 519 GGTCTACATGCTAATGGGAACATATGGAATGCAGACAGCAGCTCAAACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTGCTGGTGAATGCTCAGCAGCACC 480
DB 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGTGCTGGTGAATGCTCAGCAGCACC 638
QY 481 AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGACCCGCGAAACTTGGTCTT 540
DB 639 AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGACCCGCGAAACTTGGTCTT 698
QY 541 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
DB 699 ATCTGAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATTTACTGTGACTACAAAGTCTCAATCAGCAAGTAATTACCCCTGCTTACATGA 660
DB 759 ACCAATTTACTGTGACTACAAAGTCTCAATCAGCAAGTAATTACCCCTGCTTACATGA 818
QY 661 TAATTACCTCAAGGAAGTTCATCAATCAAGTTGAAGAAGAAACAGGCTATAAECTACTCAC 720
DB 819 TAATTACCTCAAGGAAGTTCATCAATCAAGTTGAAGAAGAAACAGGCTATAAECTACTCAC 878
QY 721 AACTGGGATGGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 780
DB 879 AACTGGAATGGGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT 938
QY 781 TTACAATACAGCAAGTAATCGTTCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 840
DB 939 TTACAATACAGCAAGTAATCGTTCCTATCCAGACGATGAATTCGAAGTGCCTTCTACCAT 998

QY 841 TGTGTGATGTTTCTAACGGTAAAGTCATTTGCCAGCTAGGACGACGCCATCAGTCAAGTAA 900
DB 999 TGTGTGATGTTTCTAACGGTAAAGTCATTTGCCAGCTAGGACGACGCCATCAGTCAAGTAA 1058
QY 901 TGTTTCCCTTCGGAATTAACCAACAGCAGTAGAACAACCCGCGACTGGGGATCAACTATGAA 960
DB 1059 TGTTTCCCTTCGGAATTAACCAACAGCAGTAGAACAACCCGCGACTGGGGATCAACTATGAA 1118
QY 961 ACCGATCACAGACTATGCTCCTCGCTGGAGTACGGTGTCTAGGATTCAACTGCTACTAT 1020
DB 1119 ACCGATCACAGACTATGCTCCTCGCTGGAGTACGGTGTCTAGGATTCAACTGCTACTAT 1178
QY 1021 CGTTCACGATGAGCCCTTATAACTA 1044
DB 1179 CGTTCACGATGAGCCCTTATAACTA 1202
RESULT 28
AX111314
LOCUS
DEFINITION Sequence 2047 from Patent WO0123604.
ACCESSION AX111314
VERSION AX111314.1 GI:13927606
KEYWORDS
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2047 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
FEATURES
Location/Qualifiers
source
1..2160
/organism="Streptococcus pneumoniae"
/strain="CS109"
/db_xref="taxon:1313"
BASE COUNT 701 a 494 c 395 g 570 t
ORIGIN
Query Match 19.8%; Score 396; DB 6; Length 2160;
Best Local Similarity 99.1%; Pred. No. 1.2e-211;
Matches 696; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TAAATCTACGACATAAATAAATCAACTCATTCGTGACTTGGGTTCTGAACGGCGCGTCAA 60
DB 159 TAAATCTACGACATAAATAAATCAACTCATTCGTGACTTGGGTTCTGAACGGCGCGTCAA 218
QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
DB 219 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACCACAGGGGATGATACCATCCGTATCCTGGGAGCTTCTTCGCGCAA 180
DB 279 TCGCTTCTTCGACCACAGGGGATGATACCATCCGTATCCTGGGAGCTTCTTCGCGCAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCACAGTGTGATTAAGTTGAC 240
DB 339 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCACAGTGTGATTAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTGTAAGGCTCAGGAACTTGGTT 300
DB 399 TTACTTTTCAACTTCGACTTCGACAGACTATTTCTGTAAGGCTCAGGAACTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
DB 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518

```

361 GGTCTACATGCTCTAATGGGAACCTATGAATGCAGACAGAGCTCAAACTACTACTATGGTAA 420
Db GGTCTACATGCTCTAATGGGAACCTATGAATGCAGACAGAGCTCAAACTACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGCTGGATGCTCAGGCACC 480
Db AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGCTGGATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTACATCCAGAACGAGCCCAAGACCGCGGAACTTGGTCTT 540
Db AAACCAATATGACCCCTATTACATCCAGAACGAGCCCAAGACCGCGGAACTTGGTCTT 698
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 600
Db ATCTGAAATGAAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAAGCAGCTCAATAC 758
QY 601 ACCAATCTATGATGGACTACAAAGTCTCAATCAGCAAGTAATTACCCTGCTTACATGGA 660
Db ACCAATCTATGATGGACTACAAAGTCTCAATCAGCAAGTAATTACCCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGAAGTCTATCAATCAAGTTGAAGAAGAAAC 702
Db TAATTACCTCAAGGAAGTCTATCAATCAAGTTGAAGAAGAAAC 860

RESULT 29
SPCS1091A SPCS1091A 2172 bp DNA linear BCT 08-SEP-1995
LOCUS S.pneumoniae CS109 pbpla gene for penicillin-binding protein la.
DEFINITION
ACCESSION 249094
VERSION 249094.1 GI:984228
KEYWORDS PBP 1A gene; penicillin-binding protein la.
SOURCE Streptococcus pneumoniae.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2172)
AUTHORS Coffey,T.J., Daniels,M., McDougal,L.K., Dowson,C.G., Tenover,F.C.
and Spratt,B.G.
TITLE Genetic analysis of clinical isolates of Streptococcus pneumoniae
with high-level resistance to expanded-spectrum cephalosporins
JOURNAL Antimicrob. Agents Chemother. 39 (6), 1306-1313 (1995)
MEDLINE 96012191
AUTHORS Spratt,B.G.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1995) Spratt B. G., University of Sussex, School
of Biology, Falmer, Brighton, East Sussex, UK, BN1 9QG
FEATURES
source
location/Qualifiers
1..2172
/organism="Streptococcus pneumoniae"
/strain="CS109 (cephalosporin-resistant clinical isolate)"
/db_xref="taxon:1313"
13..2172
/gene="pbpla"
/gene="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/citation=[1]
/codon_start=1
/transl_table=11
/evidence=experimental
/product="penicillin-binding protein la"
/protein_id="CAA88917.1"
/db_xref="GI:984229"
/db_xref="SPTREMBL:Q57114"
/translation="MNKPTILRLIKLISFLSLVIAIVLGGVFFVYVSKAPLSLSE
SKLVATSKYDNNKQIADISERRVNAQADPTDLVKAIVIEDHFRFDRGID
TTRILGAFRLNQSLOGGSLTQLOIKLTVFSTSDSTSRKAQAWLAIOLEQK
ATFOELITVYINKVYMGNGMGTAAQNYGKDLNLSLPOLALLAGMPOAPNOYDP
YSHPEAODRNLVLSEMKNOGYISAEQYKAVNPTITDGLQSLKASNYPAYMNYL
KEVINQVEETGYNLTGTGMQVYVTVNQAEQKHLMDIYNSDQIVSYFDDLDQVASTV
DVSNGKVIQAQGARHQASNVSGFTNQAVETNRDWSMKRPIPTYDGALEYGYVDSTAS
```

SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2172)
AUTHORS Spratt,B.G.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1995) Spratt B. G., University of Sussex, School
of Biology, Falmer, Brighton, East Sussex, UK, BN1 9QG
REFERENCE 2 (bases 1 to 2172)
AUTHORS Coffey,T.J., Daniels,M., McDougal,L.K., Dowson,C.G., Tenover,F.C.
and Spratt,B.G.
TITLE Genetic analysis of clinical isolates of Streptococcus pneumoniae
with high-level resistance to expanded-spectrum cephalosporins
JOURNAL Antimicrob. Agents Chemother. 39 (6), 1306-1313 (1995)
MEDLINE 96012191
FEATURES Location/Qualifiers
source 1..2172
/organism="Streptococcus pneumoniae"
/strain="CS111 (cephalosporin-resistant clinical isolate)"
/db_xref="taxon:1313"
gene 13..2172
/gene="pbpla"
CDS 13..2172
/gene="pbpla"
/standard_name="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="penicillin-binding protein 1a"
/protein_id="CAA88918.1"
/db_xref="GI:984231"
/db_xref="SPTREMBL:Q57114"
/translation="MNKPTILRLKYLISFLSLVIAIVLGGVFFVYVSKAPLSLSE
SKLVATTSKLYDNKQLIADLGSRRVNAQNDIPTDLVRAIVSIEDHFRDHRGID
TIRLQALFRLNSLQSGSTLTQQLIKLTFSTSDQTLISRKAQBAWLAIQLEQK
ATKQELIYYINKVMSNGNYGMQTAQNYGKOLNLSLPQLAAGMPQAPQYDP
YSHPEADQRLNLYSEMKNQYISAEQYKAVNPPIVDGLQSLKASNYPAYMDNYL
KEVINOVEETGYNLLTTGMDVYTNVDOEAKHLWDIYNSDQYVSPDDQLQVASTVS
DVSKGVTALGARHQASNVSEGTNOAVETNRDQSGSMKPTIDVAPALEYGVYDSTAS
IVHDVPYNGPDTPLYNWDHYVFGNITIQIALQSRNVATVELINKVLGRKATFLN
GLGIDPSMHYANAISSTNSNKYGASSEMMAAYAFANGGIYHHPKMTINKIVFS
DGSEKFSADGTRAMKETAYMTMMKTVLTGYTGRAIYLPWLPQAQKGTSTNIDE
BIERYIKNTGYVADPEMGVTRKYAMAVTGYGSRNLTPIIGDGLVAGKIVYSMTY
LSEDDQPDGMDGLYRNGEYVFKNGARSTWSSPAQPPSTESSSSSDSSTQSIN
STTPSTNNSTTTPNNTQQSNTTPDQONQPQAP"
misc_feature 1120..1122
/gene="pbpla"
/note="Active-site serine residue (Ser-370) that is
acylated by penicillins and cephalosporins"
BASE COUNT 707 a 495 c 398 g 572 t
ORIGIN
Query Match 19.8%; Score 396; DB 1; Length 2172;
Best Local Similarity 99.1%; Pred. No. 1.2e-211;
Matches 696; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TAAATCTACGACATAAAATCAACTGCTGACTGGGTTCTGAACGGCGGTCAA 60
|||||
DB 171 TAAATCTACGACATAAAATCAACTGCTGACTGGGTTCTGAACGGCGGTCAA 230
|||||
QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGCA 120
|||||
DB 231 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGCA 290
|||||
QY 121 TCGCTCTTCGACACAGGGGATTGATACCATCCGTATCCCTGGGAGCTTCTTCGCGAA 180
|||||
DB 291 TCGCTCTTCGACACAGGGGATTGATACCATCCGTATCCCTGGGAGCTTCTTCGCGAA 350
|||||
QY 181 TCTGCAAGCAATTCCTCCAGGTGGATCACTCACCACAGTTCGATTAAGTTGAC 240
|||||

Db 351 TGTGAAAGCAATTCCTCCAAAGGTGGATCAACTCAACCAACAGTTGATTAATGAC 410
QY 241 TTACTTTTCAACTTCGACTTCGACATCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 300
|||||
Db 411 TTACTTTTCAACTTCGACTTCGACATCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 470
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGCAAAATCTTGCACCTACTATATAATAA 360
|||||
Db 471 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGCAAAATCTTGCACCTACTATATAATAA 530
QY 361 GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATGGTAA 420
|||||
Db 531 GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATGGTAA 590
QY 421 AGACCTCAATAATTAAGTTTACCTCAGTTAGCTTGTGCTGGTGGATGCTCAGGCACC 480
|||||
Db 591 AGACCTCAATAATTAAGTTTACCTCAGTTAGCTTGTGCTGGTGGATGCTCAGGCACC 650
QY 481 AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGAGCCGCGGAACTTGGTCTT 540
|||||
Db 651 AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGAGCCGCGGAACTTGGTCTT 710
QY 541 ATCTGAATGAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCAGTCAATAC 600
|||||
Db 711 ATCTGAATGAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCAGTCAATAC 770
QY 601 ACCAATCTGATGGACTACAAAGTCTCAATCAGCAAGTAATACCTCTTACATGGA 660
|||||
Db 771 ACCAATCTGATGGACTACAAAGTCTCAATCAGCAAGTAATACCTCTTACATGGA 830
QY 661 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAAAC 702
|||||
Db 831 TAATTACCTCAAGGAAGTCAATCAATCAAGTTGAAGAAGAAAC 872
RESULT 31
SPN403975 489 bp DNA linear BCT 06-JAN-2001
LOCUS Streptococcus pneumoniae partial pbpla gene for penicillin binding
protein 1a, strain 960036.
DEFINITION
ACCESSION AJ403975.1 GI:12055852
VERSION pbpla gene; penicillin binding protein 1a.
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 489)
AUTHORS Overweg,K., Bogaert,D., Sluijter,M., de Groot,R. and Hermans,P.W.M.
TITLE Molecular characterization of Streptococcus pneumoniae
penicillin-resistance in the Netherlands
JOURNAL Unpublished
AUTHORS Hermans,P.W.M.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2000) Hermans P.W.M., Laboratory of Pediatrics,
Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam,
NETHERLANDS
FEATURES Location/Qualifiers
source 1..489
/organism="Streptococcus pneumoniae"
/strain="960036"
/db_xref="taxon:1313"
gene 1..489
/gene="pbpla"
CDS <1..>489
/gene="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=11
/product="penicillin binding protein 1a"
/protein_id="CAC20957.1"
/db_xref="GI:12055853"

/translation="SNYPAYMDNLYKEVINOVEETGYNLLTTGMDVYTNVDQEAQKH
LWDIYNTDEYVAYPDDELQVASTIVDSNGKVIAGLGHOSHSNVSFGINAVETNRD
WGSTMKPIITDAPALEYGYVDSTATIVHDEPNYPGTNTPTVYNWDRGYFGNITLQYAL
QOS"
BASE COUNT 154 a 117 c 98 g 120 t
ORIGIN

Query Match 19.4%; Score 387; DB 1; Length 489;
Best Local Similarity 99.6%; Pred. No. 1.4e-206;
Matches 487; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 638 AGTAATTACCCCTGCTTACATGATTAATACCTCAAGGAAGTCATCAATCAAGTTGAAGAA 697
|||||
Db 1 AGTAATTACCCCTGCTTACATGATTAATACCTCAAGGAAGTCATCAATCAAGTTGAAGAA 60
QY 698 GAAACAGGCTATAACCTACTACAACTGGGATGGATGCTTACACAAATGTAGACCAAGAA 757
|||||
Db 61 GAAACAGGCTATAACCTGCTCAACTGGGATGGATGCTTACACAAATGTAGACCAAGAA 120
QY 758 GCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATACGTTGCCCTATCCAGACGAT 817
|||||
Db 121 GCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATACGTTGCCCTATCCAGACGAT 180
QY 818 GAATTCGAAGTCGCTTCTACCATGTTGATGTTCTTAACCGGTAAGTCAATGCCCCAGCTA 877
|||||
Db 181 GAATTCGAAGTCGCTTCTACCATGTTGATGTTCTTAACCGGTAAGTCAATGCCCCAGCTA 240
QY 878 GGAGCAGCCATCAGTCAAGTAATGTTCTTCGGGAATTAACCAAGCAGTGAACCAAAAC 937
|||||
Db 241 GGAGCAGCCATCAGTCAAGTAATGTTCTTCGGGAATTAACCAAGCAGTGAACCAAAAC 300
QY 938 CGCGACTGGGGATCAACTATGAACCGATCACAGACTATGCTCCTGCTTGGAGTACGGT 997
|||||
Db 301 CGCGACTGGGGATCAACTATGAACCGATCACAGACTATGCTCCTGCTTGGAGTACGGT 360
QY 998 GTCTAGGATCAACTGCTACTATGTTTACAGTACGAGCTATGCTCCTGCTTGGAGTACGGT 1057
|||||
Db 361 GTCTAGGATCAACTGCTACTATGTTTACAGTACGAGCTATGCTCCTGCTTGGAGTACGGT 420
QY 1058 ACTCCTGTTTATAACTGGGATAGGGGCTACTTTGGCAACATCACCTTGAATACGCCCTG 1117
|||||
Db 421 ACCCTGTTTATAACTGGGATAGGGGCTACTTTGGCAACATCACCTTGAATACGCCCTG 480
QY 1118 CAACATCG 1126
|||||
Db 481 CAACATCG 489

RESULT 32
SPN403978 489 bp DNA linear BCT 06-JAN-2001
LOCUS Streptococcus pneumoniae partial.pbpla gene for penicillin binding
protein 1a, strain 950423.

ACCESSION AJ403978
VERSION 1 GI:12055858
KEYWORDS pbpla gene; penicillin binding protein 1A.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae

REFERENCE 1 (bases 1 to 489)
Overweg,K., Bogaert,D., Sluifjter,M., de Groot,R. and Hermans,P.W.M.
Molecular characterization of Streptococcus pneumoniae
penicillin-resistance in the Netherlands

Unpublished
2 (bases 1 to 489)
Hermans,P.W.M.
Direct Submission

Submitted (25-MAY-2000) Hermans P.W.M., Laboratory of Pediatrics,
Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam,
NETHERLANDS

Location/Qualifiers
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.

REFERENCE 1 (bases 1 to 489)
Overweg,K., Bogaert,D., Sluifjter,M., de Groot,R. and Hermans,P.W.M.
Molecular characterization of Streptococcus pneumoniae
penicillin-resistance in the Netherlands

Unpublished
2 (bases 1 to 489)
Hermans,P.W.M.
Direct Submission

Submitted (25-MAY-2000) Hermans P.W.M., Laboratory of Pediatrics,
Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam,
NETHERLANDS

source

1..489
/organism="Streptococcus pneumoniae"

/strain="950423"

/db_xref="taxon:1313"

1..489
/gene="pbpla"

<1..489
/gene="pbpla"

/function="peptidoglycan transglycosylase-transpeptidase"

/codon_start=1
/transl_table=11

/product="penicillin binding protein 1a"

/protein_id="CAC20960.1"

/db_xref="GI:12055858"

/translation="SNYPAYMDNLYKEVINOVEETGYNLLTTGMDVYTNVDQEAQKH
LWDIYNTDEYVAYPDDELQVASTIVDSNGKVIAGLGHOSHSNVSFGINAVETNRD
WGSTMKPIITDAPALEYGYVDSTATIVHDEPNYPGTNTPTVYNWDRGYFGNITLQYAL
QOS"

BASE COUNT 154 a 117 c 98 g 120 t
ORIGIN

Query Match 19.4%; Score 387; DB 1; Length 489;
Best Local Similarity 99.6%; Pred. No. 1.4e-206;
Matches 487; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 638 AGTAATTACCCCTGCTTACATGATTAATACCTCAAGGAAGTCATCAATCAAGTTGAAGAA 697
|||||
Db 1 AGTAATTACCCCTGCTTACATGATTAATACCTCAAGGAAGTCATCAATCAAGTTGAAGAA 60
QY 698 GAAACAGGCTATAACCTACTACAACTGGGATGGATGCTTACACAAATGTAGACCAAGAA 757
|||||
Db 61 GAAACAGGCTATAACCTGCTCAACTGGGATGGATGCTTACACAAATGTAGACCAAGAA 120
QY 758 GCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATACGTTGCCCTATCCAGACGAT 817
|||||
Db 121 GCTCAAAAACATCTGTGGGATATTTACAATACAGACGAATACGTTGCCCTATCCAGACGAT 180
QY 818 GAATTCGAAGTCGCTTCTACCATGTTGATGTTCTTAACCGGTAAGTCAATGCCCCAGCTA 877
|||||
Db 181 GAATTCGAAGTCGCTTCTACCATGTTGATGTTCTTAACCGGTAAGTCAATGCCCCAGCTA 240
QY 878 GGAGCAGCCATCAGTCAAGTAATGTTCTTCGGGAATTAACCAAGCAGTGAACCAAAAC 937
|||||
Db 241 GGAGCAGCCATCAGTCAAGTAATGTTCTTCGGGAATTAACCAAGCAGTGAACCAAAAC 300
QY 938 CGCGACTGGGGATCAACTATGAACCGATCACAGACTATGCTCCTGCTTGGAGTACGGT 997
|||||
Db 301 CGCGACTGGGGATCAACTATGAACCGATCACAGACTATGCTCCTGCTTGGAGTACGGT 360
QY 998 GTCTAGGATCAACTGCTACTATGTTTACAGTACGAGCTATGCTCCTGCTTGGAGTACGGT 1057
|||||
Db 361 GTCTAGGATCAACTGCTACTATGTTTACAGTACGAGCTATGCTCCTGCTTGGAGTACGGT 420
QY 1058 ACTCCTGTTTATAACTGGGATAGGGGCTACTTTGGCAACATCACCTTGAATACGCCCTG 1117
|||||
Db 421 ACCCTGTTTATAACTGGGATAGGGGCTACTTTGGCAACATCACCTTGAATACGCCCTG 480
QY 1118 CAACATCG 1126
|||||
Db 481 CAACATCG 489

RESULT 33

SPN403978

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SPN403978

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SPN403978

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SPN403978

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

REFERENCE	Streptococcus.	361	GGTCTACATGCTTAATGGAACTATGGAATGCAGACAGCAGCTCAAAACTACTATGGTAA	420
AUTHORS	Hakenbeck, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.	519	GGTCTACATGCTTAATGGAACTATGGAATGCAGACAGCAGCTCAAAACTACTATGGTAA	578
REFERENCE	Genetik, Imestr. 73, 1000 Berlin 33, FRG			
AUTHORS	2 (bases 1 to 2157)	421	AGACCTCAATTAATTAAGTTTACCTCAGTTAGCCCTTCTGGCTGGAATGCCTCAGGCACC	480
TITLE	Martin, C., Sibold, C. and Hakenbeck, R.			
REFERENCE	Relatedness of penicillin-binding protein la genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain	579	AGACCTCAATTAATTAAGTTTACCTCAGTTAGCCCTTCTGGCTGGAATGCCTCAGGCACC	638
JOURNAL	EMBO J. 11 (11), 3831-3836 (1992)			
MEDLINE	93010977	481	AAACCAATATGACCCCTTATTCACATCCAGAAGCAGCCCAAGACCGCGAAACTTGGTCTT	540
FEATURES	Location/Qualifiers			
source	1. .2157	639	AAACCAATATGACCCCTTATTCACATCCAGAAGCAGCCCAAGACCGCGAAACTTGGTCTT	698
	/organism="Streptococcus pneumoniae"			
	/strain="456"	541	ATTCGAAATGAAATCAAGGCTACATCTCTGTGTAACAGTATGAGAAAGCAGTCAATAC	600
	/db_xref="taxon:1313"			
	/clone="pcMK1-pcMK16"	699	ATTCGAAATGAAATCAAGGCTACATCTCTGTGTAACAGTATGAGAAAGCAGTCAATAC	758
gene	1. .2157			
	/gene="ponA"	601	ACCAATTAATGATGAGTACAAAGTCTCAATCAGCAAGTAATTTACCTGCTTACATGGA	660
	1. .>2157			
	/gene="ponA"	759	ACCAATTAATGATGAGTACAAAGTCTCAATCAGCAAGTAATTTACCTGCTTACATGGA	818
	/codon_start=1			
	/transl_table=1			
	/product="penicillin-binding protein la"			
	/protein_id="CAA48068.1"			
	/db_xref="GI:47406"			
	/translation="MNKPTILRLIKYLSIFLSLVIAAIVLGGVFFYYVSKAPSLSE			
	SKLVATSSKIYDNKNLADLGSERVAQAANDIPTDLVKAIVSIEDHRRFFDHRGID			
	SIRILGALFNLSQSGSTLQOLIKLYTSTSDTISRKAQAWLAQLQEQ			
	ATKQELIYINKVYNSNGYGMQAOYNGKDLNLSLPQLALLAGMPQAPNDYD			
	YSHPEAAQDRNLVLSMKNOGYISAEQYKAVNTPTDGLQSLKASNPAYMDNYL			
	KEVINQVEQETGYLLTGMVYTNVDOEAQKHLWDIYNSDOYVSPDDDLQVASTAT			
	DYNSKVIQAQARHOASNVSEFTNQAVETNRDWSAMKPTIDYAPAIYGVYDSTAT			
	MYNDIPYVPGTSTPYVNDRAYFGNITLOYLQQRNVAVETLNKVLGIDRAKTFIN			
	GGIDPFSHYANALSSNTESNKQYGASSERMAAAYAFANGTGYIKPMYIHKVYFS			
	DGSKFEFNVGTRAKETAYMMTDMKVTLYGTGRGAYLPWLPQAGTGYTSNYTDE			
	EVENHIKNTGVYAPDEMFGYTRKYKSMVMTGYSNRLTPIVGDGFLVAARVSRMTY			
	LSIEDTHPTDWPGLFRNGEYFKNGARPIWTEPSTQSSAESSSSSSSSSSSSSQSS			
	STTPSTNNSTTTNNNTQOSNTTDPQONQPPAQ"			
BASE COUNT	695 a 494 c 403 g 565 t			
ORIGIN				
Query Match	15.9%; Score 318; DB 1; Length 2157;			
Best Local Similarity	99.0%; Pred. No. 1.3e-167;			
Matches	668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;			
QY	1 TAAATCTAGCACAATAAATAACATCAATGCTGACTTGGGTCTGAACGCCGCGTCAA	60		
Db	159 TAAATCTAGCACAATAAATAACATCAATGCTGACTTGGGTCTGAACGCCGCGTCAA	218		
QY	61 TGCCCAAGCTAATGATATCCCAACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	120		
Db	219 TGCCCAAGCTAATGATATCCCAACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	278		
QY	121 TCCTGCTTCTCGACACAGGGGATGATACCATCCGATATCCTGGAGCTTCTTGGCCAA	180		
Db	279 TCCTGCTTCTCGACACAGGGGATGATACCATCCGATATCCTGGAGCTTCTTGGCCAA	338		
QY	181 TCCTGCAAGCAATTCCTCCCAAGGTGATCAACTCTCACCACACAGTTGATTAAAGTTGAC	240		
Db	339 TCCTGCAAGTAATTCCTCCCAAGGTGATCAACTCTCACCACACAGTTGATTAAAGTTGAC	398		
QY	241 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	300		
Db	399 TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT	458		
QY	301 AGCGATTCACTTAGACAAAAACCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360		
Db	459 AGCGATTCACTTAGACAAAAACCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	518		
LOCUS	SPONAL 2157 bp DNA linear BCT 11-MAY-1995			
DEFINITION	S.pneumoniae (681) ponA gene for penicillin-binding protein la.			
ACCESSION	X67866			
VERSION	X67866.1 GI:47407			
KEYWORDS	penicillin-binding protein la; ponA gene.			
SOURCE	Streptococcus pneumoniae.			
ORGANISM	Streptococcus pneumoniae			
REFERENCE	Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.			
AUTHORS	1 (bases 1 to 2157)			
TITLE	Hakenbeck, R.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.			
AUTHORS	Genetik, Imestr. 73, 1000 Berlin 33, FRG			
TITLE	2 (bases 1 to 2157)			
JOURNAL	Martin, C., Sibold, C. and Hakenbeck, R.			
MEDLINE	Relatedness of penicillin-binding protein la genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain			
FEATURES	EMBO J. 11 (11), 3831-3836 (1992)			
source	93010977			
Location/Qualifiers				
1. .2157				
/organism="Streptococcus pneumoniae"				
/strain="456"				
/db_xref="taxon:1313"				
/clone="pcMK1-pcMK16"				
1. .2157				
/gene="ponA"				
1. .>2157				
/gene="ponA"				
/codon_start=1				
/transl_table=1				
/product="penicillin-binding protein la"				
/protein_id="CAA48068.1"				
/db_xref="GI:47406"				
/translation="MNKPTILRLIKYLSIFLSLVIAAIVLGGVFFYYVSKAPSLSE				
SKLVATSSKIYDNKNLADLGSERVAQAANDIPTDLVKAIVSIEDHRRFFDHRGID				
SIRILGALFNLSQSGSTLQOLIKLYTSTSDTISRKAQAWLAQLQEQ				
ATKQELIYINKVYNSNGYGMQAOYNGKDLNLSLPQLALLAGMPQAPNDYD				
YSHPEAAQDRNLVLSMKNOGYISAEQYKAVNTPTDGLQSLKASNPAYMDNYL				
KEVINQVEQETGYLLTGMVYTNVDOEAQKHLWDIYNSDOYVSPDDDLQVASTAT				
DYNSKVIQAQARHOASNVSEFTNQAVETNRDWSAMKPTIDYAPAIYGVYDSTAT				
MYNDIPYVPGTSTPYVNDRAYFGNITLOYLQQRNVAVETLNKVLGIDRAKTFIN				
GGIDPFSHYANALSSNTESNKQYGASSERMAAAYAFANGTGYIKPMYIHKVYFS				
DGSKFEFNVGTRAKETAYMMTDMKVTLYGTGRGAYLPWLPQAGTGYTSNYTDE				
EVENHIKNTGVYAPDEMFGYTRKYKSMVMTGYSNRLTPIVGDGFLVAARVSRMTY				
LSIEDTHPTDWPGLFRNGEYFKNGARPIWTEPSTQSSAESSSSSSSSSSSSSQSS				
STTPSTNNSTTTNNNTQOSNTTDPQONQPPAQ"				

QY 361 GGTCTACATGCTAATGGGAACATATGAATGCAGACAGCAGCTCAAAACTACTATATGGTAA 420
|||||
Db 519 GGTCTACATGCTAATGGGAACATATGAATGCAGACAGCAGCTCAAAACTACTATATGGTAA 578
QY 421 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGCACC 480
|||||
Db 579 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCCGAAACTTGTCTTT 540
|||||
Db 639 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCCGAAACTTGTCTTT 698
QY 541 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 600
|||||
Db 699 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAAATACCTGCTTACATGGA 660
|||||
Db 759 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAAATACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
|||||
Db 819 TAATTACCTCAAGGA 833

RESULT 36
SPON44
LOCUS
DEFINITION S.pneumoniae (56742) ponA gene for penicillin-binding protein 1a.
ACCESSION X67869
VERSION X67869.1 GI:47411
KEYWORDS penicillin-binding protein 1a; ponA gene.
SOURCE Streptococcus pneumoniae.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2157)
Hakenbeck, R.
Direct Submission
TITLE Submitted (24-JUL-1992) R. Hakenbeck, Max-Planck Inst. fuer Molek.
JOURNAL Genetika, Innestr. 73, 1000 Berlin 33, FRG
REFERENCE 2 (bases 1 to 2157)
Martin, C., Sibold, C. and Hakenbeck, R.
TITLE Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain
JOURNAL EMBO J. 11 (11), 3831-3836 (1992)
MEDLINE 93010977
FEATURES Location/Qualifiers
1..2157
/organism="Streptococcus pneumoniae"
/strain="56742"
/db_xref="taxon:1313"
/clone="pcMK1-pcMK16"
1..2157
/gene="ponA"
1..>2157
/gene="ponA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="CAA48069.1"
/db_xref="GI:47412"
/db_xref="SPTREMBL:Q54949"
/translation="MNKPTILRLIKYLSFLVIAAIVLGGGVFFYVSKAPSLSE
SKLVATTSKIVDNKNOLIADLGSERRVNAQNDIPTDLVKAIVSIEDHRRFDHRCID
SIRILGAEFLRNQSNLSGGSTLTQOLIKLYFSTSDQSLSRKAQANLAIQEQK
ATKQELIYINKVTVMNSNGYMQTAQNYGKDLNLSLPALLGMPQAPNPQDP
YSHPEAQDRNRNLVSEMNQGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLY
KEVINQVQETGYNLTGMDVYTNVDOEAKHLWDIYNSDQYVSYPDLEQVASTVY
DVNSKGVIAOLGARHOASNVSGFTNQAVETNRDGSAMKPIIDYAPALEYGYDSTAT
MVNDIPYNPGTSTPYVNDRAYFGNITLQALQQSRNVAVETLNKVLQDLRAKFTLN
GLGIDYPSMHYANAISNTTESNKQYGASSBKMAAAYAAAFANGGTYGKPYIHKVYFS

BASE COUNT 696 a 494 c 403 g 564 t
ORIGIN
DGSKEFSNVGTRAMKETTAYMMTDMKTVLTYGTGRGAYLPWLPOAKTGTSNVTDE
EVENHKNITGVYADPEFVGRYKYSMAVWTGYSNRLTPIVGDGFLVAKYVRSMITY
LSEDPEDWTPMDGLFRNGEVEFKNGRPIWTEPSTQSSSTAESSSSDSSTSQSS
STTPSTNNSTTTPNNNTQOSNTTTPDOQNPNPPAP"

Query Match 15.9%; Score 318; DB 1; Length 2157;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAAATCTAGGACAATAAAACAACACTCATGTGCTGACTTGGGTTCTGAAGCGCGCTCAA 60
|||||
Db 159 TAAATCTAGGACAATAAAACAACACTCATGTGCTGACTTGGGTTCTGAAGCGCGCTCAA 218
QY 61 TGCCCAAGCTAATGATATATCCACACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
|||||
Db 219 TGCCCAAGCTAATGATATATCCACACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACCAACAGGGGGATTGATACCATCCGTATCCTGGGAGCTTCTTTCGCGCAA 180
|||||
Db 279 TCGCTTCTTCGACCAACAGGGGGATTGATTCCTATCCGTATCCTGGGAGCTTCTTTCGCGCAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTACCCCAACAGTTGATTAAGTTGAC 240
|||||
Db 339 TCTGCAAGCAATTCCTCCCAAGGTGGATCAACTCTACCCCAACAGTTGATTAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCGAGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 300
|||||
Db 399 TTACTTTTCAACTTCGACTTCGAGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTGGTT 458
QY 301 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA 360
|||||
Db 459 AGCGATTCAAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA 518
QY 361 GGTCTACATGCTAATGGGAACATATGGAATGCAGACAGCAGCTCAAAACTACTATATGGTAA 420
|||||
Db 519 GGTCTACATGCTAATGGGAACATATGGAATGCAGACAGCAGCTCAAAACTACTATATGGTAA 578
QY 421 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGCACC 480
|||||
Db 579 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGGTAATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCCGAAACTTGTCTTT 540
|||||
Db 639 AAACCAATATGACCCCTATTACATCCAGAAGCAGCCCAAGACCGCCGAAACTTGTCTTT 698
QY 541 ATCTGAATGAATAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 600
|||||
Db 699 ATCTGAATGAATAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAAATACCTGCTTACATGGA 660
|||||
Db 759 ACCAATTACTGATGGACTACAAGTCTCAAACTCAGCAAGTAAATACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
|||||
Db 819 TAATTACCTCAAGGA 833

RESULT 37
AX111407
LOCUS AX111407
DEFINITION Sequence 2140 from Patent WO0123604.
ACCESSION AX111407
VERSION AX111407.1 GI:13927699
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2157)

AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
 Picard,F.J. and Roy,P.H.
 TITLE Highly conserved genes and their use to generate probes and primers
 for detection of microorganisms
 JOURNAL Patent: WO 0123604-A 2140 05-APR-2001;
 Infectio Diagnostic (I.D.I.) INC. (CA)
 FEATURES Location/Qualifiers
 source 1..2157
 /organism="Streptococcus pneumoniae"
 /strain="670"
 /db_xref="taxon:1313"
 BASE COUNT 696 a 494 c 405 g 562 t
 ORIGIN

 Query Match 15.9%; Score 318; DB 6; Length 2157;
 Best Local Similarity 99.0%; Pred. No. 1.3e-167;
 Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

 QY 1 TAAATCTACGACAAATAAATCAACATTCGCTGACTTGGGTTCTGACGCCCGCTCAA 60
 DB 159 TAAATCTACGACAAATAAATCAACATTCGCTGACTTGGGTTCTGACGCCCGCTCAA 218

 QY 61 TGCCCAAGCTAATGATATTCACAGAGATTTGGTTAAGGCAATCGTTCTATCGAAGACCA 120
 DB 219 TGCCCAAGCTAATGATATTCACAGAGATTTGGTTAAGGCAATCGTTCTATCGAAGACCA 278

 QY 121 TCGTCTTCGACACACAGGGGGATGTACCATCGGTATCTCTGGAGGTTTCTTTGCGCAA 180
 DB 279 TCGTCTTCGACACACAGGGGGATGTATCCATCCGTATCTCTGGAGGTTTCTTTGCGCAA 338

 QY 181 TCTCGAAAGCAATTCCTCCCAAGGTGGATCAACTCTCACCCACACAGTTGATTAAGTTGAC 240
 DB 339 TCTCGAAAGTAATTCCTCCCAAGGTGGATCAACTCTCACCCACACAGTTGATTAAGTTGAC 398

 QY 241 TTACTTTTCAACTTCGACTTCCGACACAGATTTTCGTAAGGCTCAGGAAGCTTGGTT 300
 DB 399 TTACTTTTCAACTTCGACTTCCGACACAGATTTTCGTAAGGCTCAGGAAGCTTGGTT 458

 QY 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAATCTTGACCTACTATATAATAAA 360
 DB 459 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAATCTTGACCTACTATATAATAAA 518

 QY 361 GGTCTACATGCTTAATGGGAACTATGGAATCGACAGCAGCTCAAACTACTATGTTAA 420
 DB 519 GGTCTACATGCTTAATGGCAACTATGGAATCGACAGCAGCTCAAACTACTATGTTAA 578

 QY 421 AGCCTCAATAATTTAAGTTTACCTCAGTTAGCCCTGCTGCTGGAATGCTCAGGCACC 480
 DB 579 AGCCTCAATAATTTAAGTTTACCTCAGTTAGCCCTGCTGCTGGAATGCTCAGGCACC 638

 QY 481 AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACGCCGGAACCTTGGTCTT 540
 DB 639 AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACGCCGGAACCTTGGTCTT 698

 QY 541 ATCTGAAATGAAAATCAAGCTACATCTCTGCTGACAGTATCGAAGACAGTCAATAC 600
 DB 699 ATCTGAAATGAAAATCAAGGTTACATCTCTGCTGAAAGTATGAGAAGACAGTCAATAC 758

 QY 601 ACCAATTACTGATGGAGCTACAAAGTCTCAAAATCAGCAAGTAATTTACCTGCTTACATGGA 660
 DB 759 ACCAATTACTGATGGAGCTACAAAGTCTCAAAATCAGCAAGTAATTTACCTGCTTACATGGA 818

 QY 661 TAAATTACTCTCAAGGA 675
 DB 819 TAAATTACTCTCAAGGA 833

AF139883	AF139883.1	GI:5410456	
VERSION	AF139883.1	GI:5410456	
KEYWORDS			
SOURCE	Streptococcus pneumoniae.		
ORGANISM	Streptococcus pneumoniae		
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.		
AUTHORS	1 (bases 1 to 2160)		
TITLE	Coffey, T.J., Daniels, M., Enright, M.C. and Spratt, B.G. Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of Streptococcus pneumoniae arose by large recombinational replacements of the cpsA-pbpIIa region		
JOURNAL	Microbiology 145 (Pt 8), 2023-2031 (1999)		
MEDLINE	99392464		
PUBMED	10463168		
REFERENCE	2 (bases 1 to 2160)		
AUTHORS	Coffey, T.J., Daniels, M., Enright, M.C. and Spratt, B.G. Direct Submission		
TITLE	Submitted (31-MAR-1999) The Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK		
JOURNAL			
FEATURES	Location/Qualifiers		
Source	1..2160		
	/organism="Streptococcus pneumoniae"		
	/strain="Sp-865"		
	/serotype="9V"		
	/db_xref="taxon:1313"		
gene	/notes="Spanish member of the major penicillin-resistant Spanish serotype 9V clone"		
	1..2160		
	/gene="pbpIIa"		
CDS	1..2160		
	/gene="pbpIIa"		
	/function="peptidoglycan transglycosylase-transpeptidase"		
	/codon_start=1		
	/transl_table=11		
	/product="penicillin-binding protein 1a"		
	/protein_id="A043066.1"		
	/db_xref="GI:5410457"		
	/translation="MNKPTILRLKYLISFLSLVIAIIVGGGVFFVYVSKAPLSLE SKIVATNTSSKIVDNKNQILADGSSRRVNAQNDIPTDLVKAIVSTEDHRFFDHRGID SIRILGAFIRLNQSLNLOGSTLTQOLIKLTVFSTSDOTISRAQAEAWLAIOLEQK ATKQETLTVYINKVYMSNGYNGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDP YSHPEAQRRLNLILESMKNQGYISAEQYEKAVNPTIDGLSLSASNYPAYMNYL KEYINGVEQETGNLTGMDVINYVDQAKHLWDIYNSDQIVVSTPDDQLQVASTVV DVNSGVIQAQLGACASNVSTGNOAVETNRDWSAMKPTIDYAPALEYGYVSTAT MVNDIPYNTPGTSTPYNWRDAYFGNITLQALQOSRNVTAETLNKVLGDRKFTLN GLSKDPSPMHYANIASNTTSENKOYGASSERMAAAYAAFANGTGYKPMYIHKVYVFS DGSKGFEFSNVGTPAMKETTYAYMNTDMKTVLTGTCRGAFLPWLQAGKTGTSNVTDE EVENHJNTGYVAPDPMFVGYRKYYSMAVTGYSNRLPTIVGDFELVAAKYVRSMTY ESDTPHEDMTMPDGLFRNGEYFVRKNGARPIWTEFTSQSSSTAESSSSDSTSTSQSS LSTPDSNNWNTNNNTNNQNTPTDQONQNPQAP"		
BASE COUNT	698 a 494 c 403 g 565 t		
ORIGIN			
Query Match	15.9%; Score 318; DB 1; Length 2160;		
Best Local Similarity	99.0%; Pred. No. 1.3e-167;		
Matches	668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;		
Qy	1 TAAATCTACGACAAATAAATCAACTCATTCGTGACTTCGGTTCCTGAAGCCGCGGTCAA 60		
Db	159 TAAATCTACGACAAATAAATCAACTCATTCGTGACTTCGGTTCCTGAAGCCGCGGTCAA 218		
Qy	61 TGCCCAAGCTAATGATATTCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120		
Db	219 TGCCCAAGCTAATGATATTCACACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278		
Qy	121 TCGTCTTCGACACAGGGGATGTATACATCCGATCTCTGGAGCTTTCTTGGCGAA 180		
Db	279 TCGTCTTCGACACAGGGGATGTATCCATCCGATCTCTGGAGCTTTCTTGGCGAA 338		

QY	181	TCGCAAGCAATTCCTCCAGGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	240
Db	339	TCGCAAAATTAATTCCTCCAGGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	398
QY	241	TTACTTTTCAACTTCGACTTCCGACGACAGACTATTTCGTAAGGCTCAGGAAGCTTGGTT	300
Db	399	TTACTTTTCAACTTCGACTTCCGACGACAGACTATTTCGTAAGGCTCAGGAAGCTTGGTT	458
QY	301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAATAA	360
Db	459	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAAATAA	518
QY	361	GCTCTACATGCTTAATGGGAACATATGGAATGCAGACAGCAGCTCAAACTACTATGTTAA	420
Db	519	GCTCTACATGCTTAATGGCAACTATGGAATGCAGACAGCAGCTCAAACTACTATGTTAA	578
QY	421	AGACCTCAATAAATTAAGTTTACCTCAGTTAGCCTTGTGCTGGCTGGAATGCCTCAGGCACC	480
Db	579	AGACCTCAATAAATTAAGTTTACCTCAGTTAGCCTTGTGCTGGCTGGAATGCCTCAGGCACC	638
QY	481	AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCGCGAAACTTGGTCTT	540
Db	639	AAACCAATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCGCGCGAAACTTGGTCTT	698
QY	541	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	600
Db	699	ATCTGAAATGAAAAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	758
QY	601	ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATACCCTGCTTACATGGA	660
Db	759	ACCAATTTACTGATGGACTACAAAGTCTCAATCAGCAAGTAAATACCCTGCTTACATGGA	818
QY	661	TAATTTACCTCAAGGA	675
Db	819	TAATTTACCTCAAGGA	833
RESULT 39			
AFI39884			
LOCUS	AFI39884 2160 bp DNA linear BCT 26-AUG-1999		
DEFINITION	Streptococcus pneumoniae strain PO-273 penicillin-binding protein		
ACCESSION	AFI39884		
VERSION	AFI39884.1 GI:5410458		
KEYWORDS	.		
SOURCE	Streptococcus pneumoniae.		
ORGANISM	Streptococcus pneumoniae		
REFERENCE	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.		
AUTHORS	1 (bases 1 to 2160)		
TITLE	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G. Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of Streptococcus pneumoniae arose by large recombinational replacements of the cpsA-bpIIa region		
JOURNAL	Microbiology 145 (Pt 8), 2023-2031 (1999)		
MEDLINE	99392464		
PUBMED	10463168		
REFERENCE	2 (bases 1 to 2160)		
AUTHORS	Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-MAR-1999) The Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK		
FEATURES			
source	Location/Qualifiers		
	1..2160		
	/organism="Streptococcus pneumoniae"		
	/strain="PO-273"		
	/serotype="14"		
	/db_xref="taxon:1313"		
	/note="Polish variant of the major penicillin-resistant Spanish serotype 9V clone"		
gene	1..2160		

/gene="pbpla"			
1..2160			
/gene="pbpla"			
/function="peptidoglycan transglycosylase-transpeptidase"			
/codon_start=1			
/transl_table=11			
/product="penicillin-binding protein 1a"			
/protein_id="AAD43067.1"			
/translation="MNKPTILRLIKYLSIFLSLVAIAVLGGGVFFYYVSKAPSLSE			
SKLVATSSKIYDNKNOLIALGSERRVNAOANDIPTDLKAIYVIEDHREFDHRGID			
SIRLGAFLNLSQNSLOGGSTLQOLIKLYFSTSDPTISRKAQEAWLAIQLEOK			
ATKEILTYIYNKYNMNGNTQMOTAAQNYIKDLNLSLQALLAGMPQAPQNDP			
YSHPEAQDRNLVLSEMKNGYISAEQYKAVNTPIIDGLQSLKSNASNPAYMDNYL			
KEYINQVEQETGYNLLTGMVDYNNVQEAQHLWDIYNSDYYSYPDDDLQVASTV			
DVSNGKVIQALGRHQASVFGTNAVETNRDWGSAMKPIITDPALEYGVYSTAT			
MVNDIPYNPCTSTPYNMDRAYFNTILOYAQSRNVTAVENTLNVGLDRAKTFLN			
GLGIDYPSMHYANAISNNTESNKQYGASSEKMAAAYAAFAFGTYKPKMVIHKVFS			
DGSKEFSNVGTRAMKETATYMTDMKTVLTYGTGRGAYLPWLPQAGKTGTSNVTDE			
EVENHIKNTGYVADEMPFYTRYKRYSMAVWTGYSNRLTPIVGDGFLVAAKYTSMMTY			
LSEGSNPEDMNIPEGLYRNGEFVKNGARSTWSPAPQPPSTESSSSDSSTSQSS			
STTSTNNSTTTNPNNNTQQSNTTPDQONQNPQAPQ"			
BASE COUNT	693 a	494 c	407 g 566 t
ORIGIN			
Query Match	15.9%	Score 318;	DB 1; Length 2160;
Best Local Similarity	99.0%;	Pred. No. 1.3e-167;	
Matches 668;	Conservative	0; Mismatches 7;	Indels 0; Gaps 0;
QY	1	TAATAATCTACGACAATAAAATCAACTCATTCGCTGACTTGGGTTCTGAACGCGCGTCAA	60
Db	159	TAATAATCTACGACAATAAAATCAACTCATTCGCTGACTTGGGTTCTGAACGCGCGTCAA	218
QY	61	TGCCCAAGCTAATGATATCCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACCA	120
Db	219	TGCCCAAGCTAATGATATCCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACCA	278
QY	121	TGCGTTCTTCGACCACAGGGGATTCATACCATCGTATCTCTGGAGCTTTCTTGGCAA	180
Db	279	TGCGTTCTTCGACCACAGGGGATTCATTCATCGTATCTCTGGAGCTTTCTTGGCAA	338
QY	181	TCTGCAAAAGCAATTCCTCCAAAGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	240
Db	339	TCTGCAAAAGCAATTCCTCCAAAGTGGATCAACTCTCACCCAAACAGTTGATTAAGTTGAC	398
QY	241	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT	300
Db	399	TTACTTTTCAACTTCGACTTCGACACAGACTATTTCTCGTAGGCTCAGGAAGCTTGGTT	458
QY	301	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA	360
Db	459	AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTTGACCTACTATATAAATAA	518
QY	361	GCTCTACATGCTCTAATGGGAACACTATGGAATGCAGACAGCTCAAACTACTATATGTTAA	420
Db	519	GCTCTACATGCTCTAATGGCAACTATGGAATGCAGACAGCTCAAACTACTATATGTTAA	578
QY	421	AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGTGCTGGCTGGAATGCCFCAGGCACC	480
Db	579	AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGTGCTGGCTGGAATGCCFCAGGCACC	638
QY	481	AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGACCCCGGAAACTTGGTCTT	540
Db	639	AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGACCCCGGAAACTTGGTCTT	698
QY	541	ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	600
Db	699	ATCTGAAATGAAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAGCAGTCAATAC	758
QY	601	ACCAATTTACTGATGACTACAAAGTCTCAAAATCAGCAAGTAAATACCCTGCTTACATGGA	660
Db	759	ACCAATTTACTGATGACTACAAAGTCTCAAAATCAGCAAGTAAATACCCTGCTTACATGGA	818


```
FEATURES
source
    University of Oxford, South Parks Road, Oxford, Oxfordshire OX1
    3PS, UK
    Location/Qualifiers
      1..2160
        /organism="Streptococcus pneumoniae"
        /strain="PO-342"
        /serotype="14"
        /db_xref="taxon:1313"
        /note="Polish variant of the major penicillin-resistant
        Spanish serotype 9V clone"
      1..2160
        /gene="pbpla"
      1..2160
        /gene="pbpla"
        /function="peptidoglycan transglycosylase-transpeptidase"
        /codon_start=1
        /transl_table=11
        /product="penicillin-binding protein 1a"
        /protein_id="A043069.1"
        /db_xref="GI:5410463"
        /translation="MNKPTILRLIKYLSIFLSLVIAAIVLGGVFFYYVSKAPSLSE
        SKLVATSSKIYDNKNQLIADLSGSRVNAQANDIPTDLVKAIVSTEDHRFDHIGID
        SIRILGAFNLQNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQEAWLATLEQK
        ATKQELITYYINKVYMSNGYGMQTAQNYGKDLNLSLPQALLAGNPQPNQYDP
        YSHPEAQDRRLVLSEMNQGTISAEQYKAVNPITDGLQSLKASNPATMDNLY
        KEVINOQETGYNLLTTGMDVYTNVQEAQHLWDIYNSDYVSPDDQLQVSTV
        DYSNGKVIAGLARHOASNVSEFTNQAVETNRDWSAMKPTIDYAPAEIYVDSTAT
        MYNDIPYNYPGTSTPYNWDRAVFGNITLOYAQQSRNVAVETLNKVLDRKTFLN
        GLGIDYPSMHYANAISSTNTESNKQYGASSEKMAAYAAAFANGTGYKPMYTHKVYFS
        DSKKEFSNVGTRANKETAYMMTDMKTVLVIYGYGRGAYLPWLQAGTKGTSNYTDE
        EVENHIKNTGYVAPDEMFGYTRKYSMAVNTGYSNRLTPIVGDFLVAAKVYRSMYTY
        LSEGNKPTGYVNGEYFKNGARSTWSSPAQPPPTSESSSSSDSSSTSSQSS
        STPTSTNNSTTTNNPNTTQOSNTTDPQONQNPQAP"
BASE COUNT      693 a 494 c 407 g 566 t
ORIGIN

Query Match      15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATAAATAAATCAATCTGCTGACTTGGGTCTGAAGCCGCGTCAA 60
Db 159 TAAATCTAGCAATAAATAAATCAATCTGCTGACTTGGGTCTGAAGCCGCGTCAA 218
QY 61 TGCCCAAGCTAATGATATCCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGCCCAAGCTAATGATATCCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTCGACACAGGGGGATTGATACCATCCGATCCGTTATCGGAGCTTCTTGGCGAA 180
Db 279 TCGCTTCTCGACACAGGGGGATTGATTCATCCGATCCGTTATCGGAGCTTCTTGGCGAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGTGATCAACTCTCACCACACAGTTGATTAAAGTTGAC 240
Db 339 TCTGCAAGCAATTCCTCCCAAGTGATCAACTCTCACCACACAGTTGATTAAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGACAAAAGAACCAAGCAAGAAATCTTGACCTACATATAATAA 360
Db 459 AGCGATTTCAGTTAGACAAAAGAACCAAGCAAGAAATCTTGACCTACATATAATAA 518
QY 361 GGTCTACATCTCTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATGGTAA 420
Db 519 GGTCTACATCTCTAATGGGAACATATGGAATGCAGACAGAGCTCAAACTACTATGGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGATGCTCAGGCAC 480
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCTTGGCTGGGATGCTCAGGCAC 638
```

```
QY 481 AAACCAATATGACCCCTATTTCATCATCAGAACGAGCCCAAGACCGCGAAACCTTGGTCTT 540
Db 639 AAACCAATATGACCCCTATTTCATCATCAGAACGAGCCCAAGACCGCGAAACCTTGGTCTT 698
QY 541 ATCTGAATGAAAATAAAGCTACATCTCTGCTGACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAATGAAAATAAAGCTACATCTCTGCTGACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA 660
Db 759 ACCAATTACTGATGGACTACAAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA 818
QY 661 TAATTACTCTCAAGGA 675
Db 819 TAATTACTCTCAAGGA 833

RESULT 42
AF139887
LOCUS
DEFINITION
    Streptococcus pneumoniae strain M134 penicillin-binding protein 1a
    (pbpla) gene, complete cds.
ACCESSION
    AF139887
VERSION
    AF139887.1 GI:5410464
KEYWORDS
    .
SOURCE
    Streptococcus pneumoniae.
ORGANISM
    Streptococcus pneumoniae
    Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
    Streptococcus.
REFERENCE
    1 (bases 1 to 2160)
    Coffey, T.J., Daniels, M., Enright, M.C. and Spratt, B.G.
    Serotype 14 variants of the Spanish penicillin-resistant serotype
    9V clone of Streptococcus pneumoniae arose by large recombinational
    replacements of the cpsA-pbpla region
    Microbiology 145 (Pt 8), 2023-2031 (1999)
    99392464
    10463168
REFERENCE
    2 (bases 1 to 2160)
    Coffey, T.J., Daniels, M., Enright, M.C. and Spratt, B.G.
    Direct Submission
    Submitted (31-MAR-1999) The Wellcome Trust Centre for the
    Epidemiology of Infectious Disease, Department of Zoology,
    University of Oxford, South Parks Road, Oxford, Oxfordshire OX1
    3PS, UK
FEATURES
    source
      1..2160
        /organism="Streptococcus pneumoniae"
        /strain="M134"
        /serotype="14"
        /db_xref="taxon:1313"
        /note="Danish variant of the major penicillin-resistant
        Spanish serotype 9V clone"
      1..2160
        /gene="pbpla"
      1..2160
        /gene="pbpla"
        /function="peptidoglycan transglycosylase-transpeptidase"
        /codon_start=1
        /transl_table=11
        /product="penicillin-binding protein 1a"
        /protein_id="A043070.1"
        /db_xref="GI:5410465"
        /translation="MNKPTILRLIKYLSIFLSLVIAAIVLGGVFFYYVSKAPSLSE
        SKLVATSSKIYDNKNQLIADLSGSRVNAQANDIPTDLVKAIVSTEDHRFDHIGID
        SIRILGAFNLQNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQEAWLATLEQK
        ATKQELITYYINKVYMSNGYGMQTAQNYGKDLNLSLPQALLAGNPQPNQYDP
        YSHPEAQDRRLVLSEMNQGTISAEQYKAVNPITDGLQSLKASNPATMDNLY
        KEVINOQETGYNLLTTGMDVYTNVQEAQHLWDIYNSDYVSPDDQLQVSTV
        DYSNGKVIAGLARHOASNVSEFTNQAVETNRDWSAMKPTIDYAPAEIYVDSTAT
        MYNDIPYNYPGTSTPYNWDRAVFGNITLOYAQQSRNVAVETLNKVLDRKTFLN
        GLGIDYPSMHYANAISSTNTESNKQYGASSEKMAAYAAAFANGTGYKPMYTHKVYFS
        DSKKEFSNVGTRANKETAYMMTDMKTVLVIYGYGRGAYLPWLQAGTKGTSNYTDE
        EVENHIKNTGYVAPDEMFGYTRKYSMAVNTGYSNRLTPIVGDFLVAAKVYRSMYTY
        LSEGNKPTGYVNGEYFKNGARSTWSSPAQPPPTSESSSSSDSSSTSSQSS
        STPTSTNNSTTTNNPNTTQOSNTTDPQONQNPQAP"
    gene
    CDS
```

BASE COUNT 696 a 495 c 402 g 567 t
 ORIGIN
 LSEDTHPEDTMDGLFRNCEFFVKNGARSTWNSPAPQPPSTSESSSSSDSSTSQS
 STPTSTNNSTTTTNNNTQSNSTTDDQONPONPOAPQ
 Query Match 15.9%; Score 318; DB 1: Length 2160;
 Best Local Similarity 99.0%; Pred. No. 1.3e-167;
 Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Qy 1 TAAATCTACGACAAATAAATCAACTCATTTGCTGACTTTGGTCTGAACGCCGCTCAA 60
 Db 159 TAAATCTACGACAAATAAATCAACTCATTTGCTGACTTTGGTCTGAACGCCGCTCAA 218
 Qy 61 TGCCCAAGCTAATGATATPCCACAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACCA 120
 Db 219 TGCCCAAGCTAATGATATPCCACAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACCA 278
 Qy 121 TCGCTTCTTCGACACACAGGGGGATTGATACCATCGGTATCTCTGGAGCTTTCTTTGCGCAA 180
 Db 279 TCGCTTCTTCGACACACAGGGGGATTGATTCATCCGTATCTCTGGAGCTTTCTTTGCGCAA 338
 Qy 181 TCTGCAAAAGCAATTCCTCCCAAGTGGATCAACTCTCACCCACAGTTGATTAAGTTGAC 240
 Db 339 TCTGCAAAAGTAAATTCCTTCCAAAGTGGATCAACTCTCACCCACAGTTGATTAAGTTGAC 398
 Qy 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTGTAAGGCTCAGGAAGCTTGTT 300
 Db 399 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTGTAAGGCTCAGGAAGCTTGTT 458
 Qy 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
 Db 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
 Qy 361 GGTCTACATGCTTAATGGGAATCTGGAATCGACAGCAGCTCAAAACTACTATGTTAA 420
 Db 519 GGTCTACATGCTTAATGGCAACTATGGAATCGACAGCAGCTCAAAACTACTATGTTAA 578
 Qy 421 AGACCTCAATAAATTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCCTCAGGCACC 480
 Db 579 AGACCTCAATAAATTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCCTCAGGCACC 638
 Qy 481 AAACCAATATGACCCCTATTTCACATCCAGAAGAGCCCAAGACCGCGAAACTTTGCTCTT 540
 Db 639 AAACCAATATGACCCCTATTTCACATCCAGAAGAGCCCAAGACCGCGAAACTTTGCTCTT 698
 Qy 541 ATCTGGAATGAAATCAAGCTACATCTCTGCTCGACAGTATGAGAAGCAGTCAATAC 600
 Db 599 ATCTGGAATGAAATCAAGGTTACATCTCTGCTGGAACAGTATGAGAAGCAGTCAATAC 758
 Qy 601 ACCAATTACTGATGGACTACAAAGCTCAAAATCAGCAAGTAATTTACCTGCTTACATGGA 660
 Db 759 ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTTACCTGCTTACATGGA 818
 Qy 661 TAATTACCTCAAGGA 675
 Db 819 TAATTACCTCAAGGA 833
 RESULT 43
 AF139888
 LOCUS
 DEFINITION Streptococcus pneumoniae strain URU-E135 penicillin-binding protein
 la (bbpla) gene, complete cds.
 ACCESSION AF139888
 VERSION AF139888.1 GI:5410466
 KEYWORDS
 SOURCE Streptococcus pneumoniae.
 ORGANISM Streptococcus pneumoniae
 Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 Streptococcus.
 REFERENCE 1 (bases 1 to 2160)
 AUTHORS Coffey,T.J.,Daniels,M., Enright,M.C. and Spratt,B.G.

TITLE
 Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of *Streptococcus pneumoniae* arose by large recombinational replacements of the cpsA-pbpIIa region
JOURNAL MEDLINE
 Microbiology 145 (Pt 8), 2023-2031 (1999)
 99392464
PUBMED
 10463168
REFERENCE
 2 (bases 1 to 2160)
 Coffey, T.J., Daniels, M., Enright, M.C. and Spratt, B.G.
AUTHORS
 Direct Submission
TITLE
 Submitted (31-MAR-1999) The Wellcome Trust Centre for the Epidemiology of Infectious Disease, Department of Zoology, University of Oxford, South Parks Road, Oxford, Oxfordshire OX1 3PS, UK
JOURNAL
FEATURES
 Location/Qualifiers
 1. 2160
 /organism="Streptococcus pneumoniae"
 /strain="URU-E135"
 /serotype="14"
 /db_xref="taxon:1313"
 /note="Uruguayan variant of the major penicillin-resistant Spanish serotype 9V clone"
 1. 2160
 /gene="pbpIIa"
 1. 2160
 /gene="pbpIIa"
 /function="peptidoglycan transglycosylase-transpeptidase"
 /codon_start=1
 /transl_table=11
 /product="penicillin-binding protein 1a"
 /protein_id="AA043071.1"
 /db_xref="GI:5410467"
 /translation="MKPTLRLIKLYLSIFLSLVIAAIVLGGVFFYVSKAPLSLE SKIVATSSKIYDNQNSLOGSTLGLSERRVNAQNDIPTDLVKAIVSDIEHRFFDHRGID SIRILGAFGLNSLOGSTLQOLIKLTVFSTSDQTSIRKAEAWLAIQLOEK ATKOEILTYIVKYNVSGNGVMQTAQNYGKDANLSLPQALLAGMPQAPNOYDP YSHPAQRDRNLVLSMKNOYISAEQYKAVNTPITDGLQSLKASNPAYMDNLL KEVINOQEQTGVLNLTIGMDVYINVDQAKHLWDIYNSDQIVSPDDDLQVASTV SVYVYIAQLGARGHASNVSFTGNQAVETNRDQWSAKPITDYAPAIIEYGVYDSTLN MVDNIPNYPGTSPVYVNDRAVFNGLITQYALQGRNVTVATETLNKVGIDRAKTLFLN DGLGDPSPMHYANAISSNTESNKQYCASSEKMAAAYAAFANGVTYKPMYTHKVVFS DGSKKFEFSNVTRAMKETTAYMMTDMKTVLTYTGCRGVALPWLPOAGKTGTSNYTDE EVENHJKNQTVYAPDEMFGVYTRKYSHAVVTGYSNRLTPIVGDGFLVAAKAVRSMITY LSTDEHPEDMTDPGLFRNGEEFVKNGASTWNSPAPQPPPTSESSSSSSSDSSTOSSS STPTFNNSTNTNNNTQNSPTPDQNNQNPQAP"

BASE COUNT 696 a 495 c 402 g 567 t
ORIGIN

Query Match 15.9%; Score 318; DB 1; Length 2160;
 Best Local Similarity 99.0%; Pred. No. 1.3e-167;
 Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAAATCTACACAAATAAATCAACTTATGCTGACTTGGGTTCTGAACCGCGGTCAA 60
 |||||
 Db 159 TAAATCTACACAAATAAATCAACTTATGCTGACTTGGGTTCTGAACCGCGGTCAA 218
 |||||

Qy 61 TGCCCAAGCTAATGATATCCCAAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACCA 120
 |||||

Db 219 TGCCCAAGCTAATGATATCCCAAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACCA 278
 |||||

Qy 121 TCGCTTCTTCGACCACAGGGGATTTGATACATCCGTATCCTGGAGCTTCTTTCGCGCAA 180
 |||||

Db 279 TCGCTTCTTCGACCACAGGGGATTTGATTCATCCGTATCCTGGAGCTTCTTTCGCGCAA 338
 |||||

Qy 181 TCTGAAAGCAATTCCTTCGAAGGTGGATCACTCTCACCCACAGTTGATTAAGTTGAC 240
 |||||

Db 339 TCTGAAAGTAATTCCTTCGAAGGTGGATCAACTCTCACCCACAGTTGATTAAGTTGAC 398
 |||||

Qy 241 TTACTTTTCACTTCGACTTCCGACCAGACTATTTCGTAAGGCTCAGGAAGCTTGGTT 300
 |||||

Db 399 TTACTTTTCACTTCGACTTCCGACCAGACTATTTCGTAAGGCTCAGGAAGCTTGGTT 458
 |||||

Qy 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
 |||||

```
|||||
Db 459 AGCGATTTCAGTTAGAACAAAGCAACCAAGGAAATCTTCACCTACTATATAATAA 518
QY 361 GGTCATCATGCTTAATGGGAACATGATGAATGCAGACAGAGCTCAAACTACTATGTTAA 420
Db 519 GGTCATCATGCTTAATGGGAACATGATGAATGCAGACAGAGCTCAAACTACTATGTTAA 578
QY 421 AGACCTCAATAATTTAAAGTTTACCTCAGTAGCTTGGCTGGCTGGAATGCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAAGTTTACCTCAGTAGCTTGGCTGGCTGGAATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGACCCGCGAAACTTGCTCTT 540
Db 639 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGACCCGCGAAACTTGCTCTT 698
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATCTATGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCTTGCTTACATGGA 660
Db 759 ACCAATCTATGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCTTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833

RESULT 44
AF139889
LOCUS
DEFINITION
Streptococcus pneumoniae strain URU-E157 penicillin-binding protein
1a (pbpla) gene, complete cds.
ACCESSION
AF139889
VERSION
AF139889.1 GI:5410468
KEYWORDS
Streptococcus pneumoniae.
SOURCE
Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
1 (bases 1 to 2160)
REFERENCE
Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
AUTHORS
Serotype 14 variants of the Spanish penicillin-resistant serotype
TITLE
9V clone of Streptococcus pneumoniae arose by large recombinational
replacements of the cpsA-pbpla region
JOURNAL
Microbiology 145 (Pt 8). 2023-2031 (1999)
MEDLINE
99392464
PUBMED
10463168
REFERENCE
2 (bases 1 to 2160)
Coffey,T.J., Daniels,M., Enright,M.C. and Spratt,B.G.
AUTHORS
Direct Submission
TITLE
Submitted (31-MAR-1999) The Wellcome Trust Centre for the
JOURNAL
Epidemiology of Infectious Disease, Department of Zoology,
University of Oxford, South Parks Road, Oxford, Oxfordshire OX1
3PS, UK
FEATURES
Location/Qualifiers
1..2160
/organism="Streptococcus pneumoniae"
/strain="URU-E157"
/serotype="14"
/db_xref="taxon:1313"
/note="Uruguayan variant of the major penicillin-resistant
Spanish serotype 9V clone"
gene
1..2160
/gene="pbpla"
CDS
1..2160
/gene="pbpla"
/function="peptidoglycan transglycosylase-transpeptidase"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1a"
/protein_id="AAD43072.1"
/db_xref="GI:5410469"
```

```
/translation="MNKPTILRIKLSISFLVIAAIVLGGGVFFYVSKAPSLSE
SKIVATTSSKIYNQKSLADLGSERRVNAQNDPTDLVKAIVSIEDHRRFDFHRCID
SIRILGAPLUNQSNLOGSSTLTQOLIKLTVFSTSTQDTSRAQAEAWLAOLEOK
ATKQELLTVYINKVYMSNGYMGTAQONYGKDLNLSLPOLALLAGAPQAPNQYDP
YSHPEAAQDRRLNLVSEKMNQGIISAEQYKAVNTPIITDGLQSLKASNPAYMDNVL
KEVINQVEQETGYNLLTTCMDVYTNVQDEAKHLWDIYNSDQTSVSPDDDLQVASTV
DYSNKGVIAOLGARHOASNVSEFQNAIVETNRDWSAMKPIITDIYPAIEYGVDSAT
MYNDIPYNPYPGTSTPYNWDRAVFGNITLQYALQOSRNVTAVETLNKVLGIDRAKTFPLN
GLGIDYPSMHYANAISNTTNSKNOYGASSERMAAAFAAFANGVTYKPMYTHKVVS
DGSKKEFSNVGTTRAMKETAYMNTDMKMTLVYGTGGRGAVLPKPOAGKTGTSNKTDE
EVENHINKTGYVAPDEMFGVYTRKYSMAVWTGYSNRLTPIVGDGELVAAKYVRSMTY
LSEDIHPEDWTMPDGLFRNGEFVFKNGARSTWNSPAPQPPPTSESSSSSDSSTSOSS
STTPSTNNSTTTNNPNNTQOSNTTDPQONPQAP"
BASE COUNT 696 a 495 c 402 g 567 t
ORIGIN
```

```
Query Match 15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-17;
Matches 668; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1 TAAATCTACGACAATAAAATCAACTCATTGCTGACTTGGTTCGAAACGCCGCGTCAA 60
Db 159 TAAATCTACGACAATAAAATCAACTCATTGCTGACTTGGTTCGAAACGCCGCGTCAA 218
QY 61 TGCCCAAGCTAATGATATTCCTCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGCCCAAGCTAATGATATTCCTCCACAGATTTGGTTAAGGCAATCGTTTCTATCGAAGACCA 278
QY 121 TCCTCTTCGACACAGGGGGATTGATACCATCCGATCCCTGGAGCTTCTTGGCGAA 180
Db 279 TCCTCTTCGACACAGGGGGATTGATACCATCCGATCCCTGGAGCTTCTTGGCGAA 338
QY 181 TCTGCAAGCAATTCCTCCCAAGTGGATCACTCACCACACAGTTGATTAAAGTTGAC 240
Db 339 TCTGCAAGTAAITCCCTCCCAAGTGGATCACTCACCACACAGTTGATTAAAGTTGAC 398
QY 241 TTACTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 TTACTTTCAACTTCGACTTCCGACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
Db 459 AGCGATTTCAGTTAGACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
QY 361 GGTCTACATGCTTAATGGGAACATGGAATGCAGACAGAGCTCAAACTACTATGTTAA 420
Db 519 GGTCTACATGCTTAATGGGAACATGGAATGCAGACAGAGCTCAAACTACTATGTTAA 578
QY 421 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGTGAATGCTCAGGCACC 480
Db 579 AGACCTCAATAATTTAAAGTTTACCTCAGTTAGCTTGGCTGGTGAATGCTCAGGCACC 638
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGACCCGCGAAACTTGCTCTT 540
Db 639 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGACCCGCGAAACTTGCTCTT 698
QY 541 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAAATGAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC 758
QY 601 ACCAATCTATGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCTTGCTTACATGGA 660
Db 759 ACCAATCTATGATGGACTACAAGTCTCAAAATCAGCAAGTAAATACCTTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833
```

```
RESULT 45
AF139890
LOCUS
AF139890 2160 bp DNA linear BCT 26-AUG-1999
```


gene 1. .2160
 CDS /gene="pbpla"
 /gene="pbpla"
 /note="peptidoglycan transglycosylase-transpeptidase"
 /codon_start=1
 /transl_table=11
 /product="penicillin-binding protein 1a"
 /protein_id="AAD48430.1"
 /db_xref="GI:5726355"
 /translation="MKNPILRLIKVLSISFLSLVIAAIVLGGVFFYVYVSPKLSLE
 SKLVATSSKIYDNKNQIADLGSERRVNAQANDIPDLVAKIATVSTEDHRFPDRGID
 SIRILAFENLOSNSIOGSGSTLTOLIKLTVESTSDQTSRKAQEWLAQLQSK
 ATKOEILTYINRVYNSNGYQMOTRAQNYIGKDLNLSLPQALLAGMPQAPNQIDP
 YSHEPAADRRIINVLSEMNQGVISAEQYEKAVNTIPITDGLQSLKASNPAYMDNYL
 KEYINQVQGTYNLTGDMVYTNVDEQAKHLWDIYNSDOYVSPDDQLQVASTV
 DVSNGKVIAGLGRHQAHSNVSGFTNQAVETNRDWSAMKPIIDYAPAEYGVYDSTAT
 MWNDIPYNGPTSTPYVNDRAVFNQITLQYALQOQRNVATVELKNKGLGDRAKTFLN
 GLGIDYPSHHYANAISNTTESNKQYGASSEKMAAAYAFANGGTYIKPMYIHKVYFS
 DGSKFEFNVGTFRAMKETATYIMDTDMKTVLITYTGRTGAFPLWLPQAGKTGTSNYTDE
 EVENHJKNITGVAPDMEFYGTSKYSMAVMTGYSNRLTPIVGDGFLVAAKVYKSMITY
 LSEDTPEWMTDGLGFNGEYFVKNGARSTWNPSAPQPPPTSESSSSSSDSSTQSS
 STTPSTNNSTTTNNPNTQOSNTTPDQONQNPQAP"

BASE COUNT 696 a 495 c 402 g 567 t
 ORIGIN

	Query Match	15.9%; Score 318; DB 1; Length 2160;
	Best Local Similarity 99.0%; Pred. No. 1.3e-167;	
	Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
Qy 1	TAAATCTACGACAAATAAATCAACTATGTGCTGGTTCGAAGCCGCGTCAA	60
Db 159	TAAATCTACGACAAATAAATCAACTATGTGCTGGTTCGAAGCCGCGTCAA	218
Qy 61	TGCCCAAGCTAATGATATCCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACCA	120
Db 219	TGCCCAAGCTAATGATATCCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACCA	278
Qy 121	TCGCTTCTTCGACCACAGGGGATTGATACCATCCGTATCCTGGGAGCTTCTTTCGCGAA	180
Db 279	TCGCTTCTTCGACCACAGGGGATTGATTCATCCGTATCCTGGGAGCTTCTTTCGCGAA	338
Qy 181	TCTGCAAGCAATTCCTCCNAGTGGATCACTCTCACCCACAGTTGATTAAGTTGAC	240
Db 339	TCTGCAAGCAATTCCTCCNAGTGGATCACTCTCACCCACAGTTGATTAAGTTGAC	398
Qy 241	TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCCTGTAAGGCTCAGGAAGCTTGGTT	300
Db 399	TTACTTTTCAACTTCGACTTCCGACACAGACTATTTCCTGTAAGGCTCAGGAAGCTTGGTT	458
Qy 301	AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360
Db 459	AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	518
Qy 361	GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGCAGCTCAAAAACCTACTATGATAA	420
Db 519	GGTCTACATGCTTAATGGGAACATATGGAATGCAGACAGCAGCTCAAAAACCTACTATGATAA	578
Qy 421	AGACCTCAATAATTTAGTTTACCTCAGTTAGCCTTCTGGCTGGAATGCTCAGGCACC	480
Db 579	AGACCTCAATAATTTAGTTTACCTCAGTTAGCCTTCTGGCTGGAATGCTCAGGCACC	638
Qy 481	AAACCAATATGACCCCTATTTCACATCCAGAGAGAGCCCAAGACCCGCGAAGCTTGGTCTT	540
Db 639	AAACCAATATGACCCCTATTTCACATCCAGAGAGAGCCCAAGACCCGCGAAGCTTGGTCTT	698
Qy 541	ATCTGGAATGAAAATCAAGGCTACATCTCTGCTGAAACAGTATGAGAAAGCAGTCAATAC	600
Db 699	ATCTGGAATGAAAATCAAGGTTACATCTCTGCTGAAACAGTATGAGAAAGCAGTCAATAC	758
Qy 601	ACCAATTACTGATGGACTACAAGCTCTCAATCAGCAAGTAATTACCTGCTTACATGGA	660

Db 759 ACCAATCTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAATTAACCTGTTACATGGA 818

Qy 661 TAATTACCTCAAGGA 675
 |||||

Db 819 TAATTACCTCAAGGA 833

RESULT 47
 AF210749
 LOCUS
 DEFINITION Streptococcus pneumoniae isolate SP 1053 penicillin-binding protein 1A (pbpla) gene, complete cds.
 ACCESSION AF210749
 VERSION AF210749.1 GI:5633344
 KEYWORDS
 SOURCE Streptococcus pneumoniae.
 ORGANISM Streptococcus pneumoniae.
 Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae; Streptococcus.

REFERENCE 1 (bases 1 to 2160)
 Ferroni,A. and Berche,P.
 Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children
 J. Med. Microbiol. 50 (9), 828-832 (2001)
 21432820
 PUBMED 11549185
 REFERENCE 2 (bases 1 to 2160)
 Ferroni,A. and Berche,P.
 Direct Submission
 Submitted (03-DEC-1999) Microbiology, Hospital
 Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France
 Location/Qualifiers
 1. .2160
 /organism="Streptococcus pneumoniae"
 /serotype="23F"
 /isolate="SP 1053"
 /db_xref="taxon:1313"
 1. .2160
 /gene="pbpla"
 1. .2160
 /gene="pbpla"
 /codon_start=1
 /translation="transl_table=1"
 /product="penicillin-binding protein 1A"
 /protein_id="AAFL17259.1"
 /db_xref="GI:5633345"
 /translation="MKPTLRLIKYLSIFLSIVTAAIVLGGVFYVYVSKAPLSLE
 SKLVATSSKIYNKNKLIADLGSERVNAQANDPTDLKAIVSIEDHRFFDHGID
 SIRILGALNLSQSGGISTITQLIKITYSTSDQTSIRKAEAWLA10LSQK
 ATKQELITYYINKVNSNGYGMQTAQNTYGRDNLNLSLPQALLAGMPQAPNQDP
 YSHPEAADRRNLVLSMKNGQVISAQVEKAVNTPTDGLQSLKASNPAYMDNVL
 KEVINOQEQTGYNLLTGMDVNTVDQEAQKHLWDIYNSDQVSYVDPDDLOVASTV
 DVSNKVIQLGARHQASNSVSGTQNAVETNRDWSAMKPTIDYPAIEGYVDSTAT
 MVNDIPYPTGTPVTVNDRAVFGNITLQYALQQSRNVATVETLNKRVGLDRKATFN
 GLGDTPSMHYANAISNTTSESKQYGASERMAAYAAFANGGIYHKPMYINKVVF
 DSGKKFEVDYGTMRKETTYAMTMMETKTVLAYGTGRGAYLPWLAGAGKTGTNTYDD
 EIKHKIKTCYVAPDEMFGYTRYKYSMAVVTGYSNRLPTIVGDGFLVAAKYYRSMITY
 LSEDTHPEDMTWMDGLFRNGEYFVKNGARSTWNSPAPQPPSTESSSSSSSDSTSSOSS
 STPTPSTNNSTTTPNNNTQSNSTTPDQONQNPQAP"

BASE COUNT 699 a 494 c 404 g 563 t
 ORIGIN

Query Match 15.9%; Score 318; DB 1; Length 2160;
 Best Local Similarity 99.0%; Pred. No. 1.3e-167;
 Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAAATCTACACAAATAAATCAACTCATGTGACTTGGTTCTGAAGCCGCGTCAA 60
 |||||

Db 159 TAAATCTACACAAATAAATCAACTCATGTGACTTGGTTCTGAAGCCGCGTCAA 218
 |||||

Qy 61 TGCCCAAGCTAATGATATTCCACAGACTTTTGGTTAAAGGCAATCGTTTCTATCGAAGACCA 120

	Query Match	15.9%	Score 318;	DB 1:	Length 2160;
	Best Local Similarity	99.0%;	Pred. No.	1.3e-167;	
	Matches 668;	Conservative	0;	Mismatches 7;	Indels 0; Gaps 0;
Oy	1	TAAATCTACGACAATAAAAAATCAACTCATTTGTGACTTTGGTTC	GAAAGCCCGCGTCAA	60	
Db	159	TAAATCTACGACAATAAAAAATCAACTCATTTGTGACTTTGGTTC	GAAAGCCCGCGTCAA	218	
Oy	61	TGCCCAAGCTAATGATATTCCACAGATTTTGGTTTAAGGCAATCGTTTCTATCGAAGACCA	AACCA	120	

[illegible]

```

/isolate="SP 1470"
/db_xref="taxon:1313"
1. 2160
/gene="pbpla"
1. 2160
/cds
/gene="pbpla"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AAFL17260.1"
/db_xref="GI:6563347"
/translation="MKPPTLRLIKYLSISFLSLVIAAIVLGGVFFYVVKRAPSLSSE
SKIVATSSKIYDNKNOLITADLSGRVNAQANDIPTDLKVAIVSDEHRRFDDHRGID
ATREILATFLNQLSNGSGSTTQOLIKLTFTYSTSQDTSRKQAEAWLAITQLBQK
ATRKELIYYINRVYMSNGSMGTAAQNYVGRDLNNLSLPQLALLAGMPQAPNMDYLV
YSEPEAQRDRNLVLSEMKNOGYISAEQYKAVNTPTDGLQSLKASNVPAVMNLYL
KEVINOVEDRNLGTGMVDVYTNVDQEAQHLWDIYNDSQDVSYSPDDLOLVASTVY
DVNSKVIQALGARHQAASNVSGTNAQVETNRDWGSAMKPTDYAPAIKGYVDSTAT
MVNDIPNYPGTSTPTVNDYRATVFNITLQYALQQSRNVTAVELTNKVGDRUKTEFLN
GLGIDYPSHMYANISNTDESNAQYKQYKASGEMAAAYAAAFANGTYYIKPMYIHKVYDE
DGSKKFFSNVTGRAMKETAYMTDMKMTVLYTGTGRGAYLPWLPAQGTGTGTSNYTDE
EVENHILCTYVADPMFVGYTRKYSMAVNTGYSNRLTPIVGDGFLVAAKVVYRSMITY
LSEDTHPEQDMTPDGLFRNGEFVKNGAPEINTEPSTQOQSSAESSSSSDSDSTSSQSS
STPTSPNNSTTTPNNNTQOSVTPQOQNQPQAP"
BASE COUNT      697 a 493 c 405 g 565 t
ORIGIN

Query Match      15.98; Score 318; DB 1; Length 2160;
Best Local Similarity 99.08; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAAATCTACGACAAATAAATCAACTCATTGCTGACTTGGGTCTCGAACGCCGCGTCAA 60
Db 159 TAAATCTACGACAAATAAATCAACTCATTGCTGACTTGGGTCTCGAACGCCGCGTCAA 218

Qy 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACCA 120
Db 219 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAAGCAATCGTTTCTATCGAAGACCA 278

Qy 121 TCGCTTCTTCGACCACAGGGGGATGTATACCCTCGCTATCCTCGGAGCTTCTTTCGCCAA 180
Db 279 TCGCTTCTTCGACCACAGGGGGATGTATTCATCCGCTATCCTCGGAGCTTCTTTCGCCAA 338

Qy 181 TCTCAAAAGCAATTCCTCCAAAGGTGGATCAACTCTCACCCACACAGTGTATTAAGTTGAC 240
Db 339 TCTCAAAAGTAAATTCCTCCAAAGGTGGATCAACTCTCACCCACACAGTGTATTAAGTTGAC 398

Qy 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
Db 399 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458

Qy 301 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAGAATACTTGTACCTACTATATAATAA 360
Db 459 AGCGATTTCAGTTAGAACAAAAGCAACCAAGCAGAATACTTGTACCTACTATATAATAA 518

Qy 361 GGTCTACATGCTTAATGGGAACATAATGAAATGCAGACAGCAGCTCAAAAACCTACTATGGTAA 420
Db 519 GGTCTACATGCTTAATGGCAACTATGGAATGCAGACAGCAGCTCAAAAACCTACTATGGTAA 578

Qy 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCCTTGCTTGGCTGGAATGCCTCAGGCACC 480
Db 579 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCCTTGCTTGGCTGGAATGCCTCAGGCACC 638

Qy 481 AAACCAATATGACCCCTATTTCACATCCAGAAAGCAGCCCAAGACCGCGGAAACCTTGGTCTTT 540
Db 639 AAACCAATATGACCCCTATTTCACATCCAGAAAGCAGCCCAAGACCGCGGAAACCTTGGTCTTT 698.

Qy 541 ATCTGAAATGAAAATCAAGCTACATCTCTGCTGAACTATGAGAAAGCAGTCAATAC 600
Db 699 ATCTGAAATGAAAATCAAGTTACATCTCTGCTGAACTATGAGAAAGCAGTCAATAC 758

Qy 601 ACCAATTACTGATGGACTACAAAGCTTCAAAATCAGCAAGTAATTACCCTGCTTACATATGGA 660

```

```
|||||
Db 759 ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833

RESULT 49
AF210751
LOCUS AF210751 2160 bp DNA linear BCT 17-SEP-2001
DEFINITION Streptococcus pneumoniae isolate SP 22861 penicillin-binding
ACCESSION protein 1A (pbpla) gene, complete cds.
VERSION AF210751
KEYWORDS AF210751.1 GI:6563348
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Ferroni,A. and Berche,P.
TITLE Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children
JOURNAL J. Med. Microbiol. 50 (9), 828-832 (2001)
MEDLINE 21432820
PUBMED 11549185
REFERENCE 2 (bases 1 to 2160)
AUTHORS Ferroni,A. and Berche,P.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1999) Microbiology, Hospital
Necker-Enfants-Malades, 149 rue de Sevres, Paris 75015, France
FEATURES
Source
1..2160
/organism="Streptococcus pneumoniae"
/serotype="23F"
/isolate="SP 22861"
/db_xref="taxon:1313"
1..2160
/gene="pbpla"
1..2160
/gene="pbpla"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 1A"
/protein_id="AA17261.1"
/db_xref="GI:6563349"
/translation="MNKPTILRLIKYLSIFSLVIAIIVLGGVFFYYVSKAPLSLSE
SKLVATTSKIYDNKQLIADLGSERRVNAQNDIPTDLVKAIVSDIEDHFRFDRHIGID
STRILGAFRLNLSQSGSTLTQOLIKLTVFSTSDOTISRKAQEAWLAIQLEQK
ATKQELITYINKVYMSNGYGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNDYDP
YSHPEAODRNLVLSMKNGYISAEQYKAVNPTITDGLSLKASNYPAYMNYL
KEVINQVQETGYNLTGMDVYTNVDQAQRHLMDIYNSDQIVSTPDDDLQVASTV
DVNSKVIYALGARHOASNVSEFTNOAVETNRDWSAMKPTIDYAPATEYGVYDSTAT
MYNDIPYIPGSTVYWNDRAYFGNITLQALQSRNVYATVETLNKVLDRAKTFLN
GLGIDYPSHHYANALSSNTESNKQYGASSEKMAAAYAFANGGIYHKPMYINKVYFS
DGSKFEPSDVGTRAMKETAYMMETMMKTVLAVGTCRGAYLPWLAQAGKTGTSNYTDD
ELEKHKITGYVAPDEMFGVTRKYSMVAVTGYSNRLPIVGDGFLVAAKGVYRSMITY
LSEGNPDNPNPEGLIRNGEVRFNKARSTWSSPAPQPPSTSSSSSSSDSSSTSQSN
STTPSTNNSTNNNTNNNTQQSNSTPDQQNQNPQAP"
BASE COUNT 700 a 491 c 405 g 564 t
ORIGIN

Query Match 15.9%; Score 318; DB 1; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1.3e-167;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAAATCTAGCAATAAATAACACTCATGCTGACTGGGTCTGAAGCGCGGTCAA 60
|||||
Db 159 TAAATCTAGCAATAAATAACACTCATGCTGACTGGGTCTGAAGCGCGGTCAA 218
```

```
QY 61 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGAATCGTTTCTATCGAAGACCA 120
|||||
Db 219 TGCCCAAGCTAATGATATATCCACAGATTGGTTAAGGAATCGTTTCTATCGAAGACCA 278
QY 121 TCGCTTCTTCGACCACAGAGGGGATTGATACCATCCGATATCCTGGGAGCTTCTTGGCAA 180
|||||
Db 279 TCGCTTCTTCGACCACAGAGGGGATTGATTCATCCGATATCCTGGGAGCTTCTTGGCAA 338
QY 181 TCTGCAAGCAATTCCCTCCAAAGTGATCAACTCTCACCAACAGTGTGATTAAAGTTGAC 240
|||||
Db 339 TCTGCAAGCAATTCCCTCCAAAGTGATCAACTCTCACCAACAGTGTGATTAAAGTTGAC 398
QY 241 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 300
|||||
Db 399 TTACTTTTCAACTTCGACTTCCGACCAGACTATTTCTCGTAAGGCTCAGGAAGCTTGGTT 458
QY 301 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 360
|||||
Db 459 AGCGATTTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA 518
QY 361 GGTCTACATGCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAACTACTATGTAA 420
|||||
Db 519 GGTCTACATGCTTAATGGGAACCTATGGAATGCAGACAGCAGCTCAAACTACTATGTAA 578
QY 421 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTCAGGCACC 480
|||||
Db 579 AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGCTGGCTGGAATGCTCAGGCACC 638
QY 481 AAACCATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCGCGGAACCTTGGTCTT 540
|||||
Db 639 AAACCATATGACCCCTATTTCACATCCAGAGCAGCCCAAGACCGCGGAACCTTGGTCTT 698
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAACAGTCAATAC 600
|||||
Db 699 ATCTGAAATGAAAAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAAACAGTCAATAC 758
QY 601 ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA 660
|||||
Db 759 ACCAATTACTGATGGACTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGA 818
QY 661 TAATTACCTCAAGGA 675
Db 819 TAATTACCTCAAGGA 833
```

```
RESULT 50
AX110436
LOCUS AX110436 2160 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 1169 from Patent WO0123604.
ACCESSION AX110436
VERSION AX110436.1 GI:13926728
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae
ORGANISM Streptococcus pneumoniae
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2160)
AUTHORS Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 1169 05-APR-2001;
Infectio Diagnostie (I.D.I.) INC. (CA)
FEATURES
source
1..2160
/organism="Streptococcus pneumoniae"
/strain="SP-665"
/db_xref="taxon:1313"
BASE COUNT 698 a 494 c 403 g 565 t
ORIGIN

Query Match 15.9%; Score 318; DB 6; Length 2160;
```

Best Local Similarity 99.0%; Pred. No. 1.3e-167;			
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;			
QY	1	TAAGTCTAGCAATAAATAAATCAACTCATTTGCTGACTTGGGTTCTGAAACGCGCGTCAA	60
Db	159	TAAGTCTAGCAATAAATAAATCAACTCATTTGCTGACTTGGGTTCTGAAACGCGCGTCAA	218
QY	61	TGCCCCAAGCTAATGATATTCCTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	120
Db	219	TGCCCCAAGCTAATGATATTCCTCCACAGATTGGTTAAGGCAATCGTTTCTATCGAAGACCA	278
QY	121	TCGCTTCTTCGACACAGGGGGATTGATACCATCCGTATCCCTGGGAGCTTCTTGGCGCAA	180
Db	279	TCGCTTCTTCGACACAGGGGGATTGATCCATCCGTATCCCTGGGAGCTTCTTGGCGCAA	338
QY	181	TCGTCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCCACAGTTGATTAAGTTGAC	240
Db	339	TCGTCAAGTAAATTCCTCCAAAGTGGATCAACTCTCACCCACAGTTGATTAAGTTGAC	398
QY	241	TTACTTTTCAACTTCGACTTCCGACGAGACTATTTCTCGTAAGGCTCAGGAAGCTTTGGTT	300
Db	399	TTACTTTTCAACTTCGACTTCCGACGAGACTATTTCTCGTAAGGCTCAGGAAGCTTTGGTT	458
QY	301	AGCGATTCACTTAGACAAACCAACCAACCAAGCAAGAACTTCACCTACTATATAATAA	360
Db	459	AGCGATTCACTTAGACAAACCAACCAACCAAGCAAGAACTTCACCTACTATATAATAA	518
QY	361	GGTCTACATCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTTACTATGGTAA	420
Db	519	GGTCTACATCTTAATGGGAACATATGGAATGCAGACAGAGCTCAAAACTTACTATGGTAA	578
QY	421	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGATGCCTCAGGCACC	480
Db	579	AGACCTCAATAATTTAAGTTTACCTCAGTTAGCCTTGGCTGGGATGCCTCAGGCACC	638
QY	481	AAACCAATATGACCCCTATTCACATCCAGACAGAGCCCGGAACTTGGTCTT	540
Db	639	AAACCAATATGACCCCTATTCACATCCAGACAGAGCCCGGAACTTGGTCTT	698
QY	541	ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600
Db	699	ATCTGAAATGAAAAATCAAGGTTACATCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	758
QY	601	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTACCTGCTTACATGGA	660
Db	759	ACCAATTACTGATGGACTACAAAGTCTCAATCAGCAAGTAATTACCTGCTTACATGGA	818
QY	661	TAATTACCTCAAGGA	675
Db	819	TAATTACCTCAAGGA	833

Search completed: June 13, 2002, 12:36:59
Job time: 14246 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 10:52:32 ; Search time 325.44 Seconds
(without alignments)
10546.067 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATATACGACATAAAA.....ATCCTCAACGACACCA 1999

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 15
Total number of hits satisfying chosen parameters: 3497

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :				
N_Geneseq_032802:*				
1:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*			
2:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*			
3:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*			
4:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*			
5:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*			
6:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*			
7:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*			
8:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*			
9:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*			
10:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*			
11:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*			
12:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*			
13:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*			
14:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*			
15:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*			
16:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*			
17:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*			
18:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*			
19:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*			
20:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*			
21:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*			
22:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*			
23:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*			
24:	/SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	19	AAV27323 Streptococcus pneu
2	1999	100.0	10711	19	AAV52278 Streptococcus pneu
3	1438	71.9	2048	17	AT08027 S. pneumoniae peni
4	1438	71.9	2160	22	AAH02055 Streptococcus pneu
5	1438	71.9	2160	23	AAAS5688 Streptococcus pneu
6	1198	59.9	1249	22	AAH02067 Streptococcus pneu
7	1014	50.7	1260	20	AAZ35952 Streptococcus pneu
8	940	47.0	1260	20	AAZ35953 Streptococcus pneu
9	903	45.2	1260	20	AAZ35950 Streptococcus pneu

10	903	45.2	1260	20	AAZ35951 Streptococcus pneu
11	838	41.9	1260	20	AAZ35939 Streptococcus pneu
12	824	41.2	1260	20	AAZ35947 Streptococcus pneu
13	796	39.8	960	16	AAQ83241 Streptococcus pneu
14	764	38.2	1260	20	AAZ35948 Streptococcus pneu
15	745	37.3	1260	20	AAZ35949 Streptococcus pneu
16	620	31.0	930	22	AAH02060 Streptococcus pneu
17	590	29.5	1199	22	AAH01023 Streptococcus pneu
18	581	29.1	1260	20	AAZ35945 Streptococcus pneu
19	567	28.4	1229	22	AAH01026 Streptococcus pneu
20	554	27.7	1260	20	AAZ35943 Streptococcus pneu
21	541	27.1	2172	19	AAZ96311 S. pneumoniae deri
22	396	19.8	2160	22	AAH02054 Streptococcus pneu
23	318	15.9	2157	22	AAH02147 Streptococcus pneu
24	318	15.9	2160	22	AAH01178 Streptococcus pneu
25	318	15.9	2160	22	AAH02056 Streptococcus pneu
26	266	13.3	1260	20	AAZ35946 Streptococcus pneu
27	228	11.4	1195	22	AAH02059 Streptococcus pneu
28	209	10.5	1260	20	AAZ35944 Streptococcus pneu
29	178	8.9	1212	22	AAH02063 Streptococcus pneu
30	178	8.9	1222	22	AAH01025 Streptococcus pneu
31	175	8.8	1220	22	AAH01022 Streptococcus pneu
32	159	8.0	1216	22	AAH02069 Streptococcus pneu
33	129	6.5	1201	22	AAH01021 Streptococcus pneu
34	127	6.4	1212	22	AAH01014 Streptococcus pneu
35	127	6.4	1223	22	AAH01017 Streptococcus pneu
36	126	6.3	1211	22	AAH01024 Streptococcus pneu
37	126	6.3	1218	22	AAH01016 Streptococcus pneu
38	122	6.1	1213	22	AAH01015 Streptococcus pneu
39	122	6.1	1214	22	AAH01018 Streptococcus pneu
40	121	6.1	1207	22	AAH01020 Streptococcus pneu
41	120	6.0	1212	22	AAH01013 Streptococcus pneu
42	115	5.8	1242	22	AAH02064 Streptococcus pneu
43	110	5.5	1223	22	AAH01019 Streptococcus pneu
44	104	5.2	1225	22	AAH02065 Streptococcus pneu
45	101	5.1	1260	20	AAZ35940 Streptococcus pneu
46	101	5.1	1260	20	AAZ35941 Streptococcus pneu
47	97	4.9	930	22	AAH02057 Streptococcus pneu
48	97	4.9	1260	20	AAZ35942 Streptococcus pneu
49	90	4.5	782	22	AAH02071 Streptococcus pneu
50	90	4.5	813	22	AAH01655 Streptococcus pneu
51	73	3.7	1225	22	AAH01027 Streptococcus pneu
52	44	2.2	930	22	AAH02058 Streptococcus pneu
53	25	1.3	25	22	AAH01226 S. pneumoniae pp1
54	25	1.3	8395	20	AAZ35918 Enterococcus faeca
55	24	1.2	24	20	AAZ35914 Streptococcus pneu
56	24	1.2	24	20	AAZ35919 Streptococcus pneu
57	23	1.2	33	19	AAZ27437 Streptococcus pneu
58	22	1.1	22	22	AAH01201 S. pneumoniae pp1
59	22	1.1	22	22	AAH02022 S. pneumoniae pp1
60	22	1.1	2337	23	AAAS2994 Enterococcus faeca
61	20	1.0	20	22	AAH01134 S. pneumoniae pp1
62	20	1.0	20	22	AAH02031 S. pneumoniae pp1
63	20	1.0	25	22	AAH01227 S. pneumoniae pp1
64	20	1.0	550	22	AAH02283 Streptococcus pneu
65	20	1.0	554	22	AAH02066 Streptococcus pneu
66	20	1.0	680	20	AAZ37006 Human cd37 nuclei
67	20	1.0	5124	21	AAZ75761 Human ORF1316
68	20	1.0	5410	22	ABA09682 Human bone marrow
69	20	1.0	5766	22	AAI57952 Human polynucleoti
70	20	1.0	7268	22	ABA09598 Human bone marrow
71	19	1.0	19	19	AAV64673 S. pneumoniae pp2
72	19	1.0	19	19	AAV64672 S. pneumoniae pp2
73	19	1.0	428	21	AAAC38165 Arabidopsis thalia
74	19	1.0	822	23	AAAS53333 Haemophilus influe
75	19	1.0	1295	21	AAAC51778 Arabidopsis thalia
76	19	1.0	2193	19	AAV19373 S. pneumoniae peni
77	19	1.0	5448	19	AAV65229 DNA encoding a S.
78	19	1.0	23041	22	AAAS26779 Human genomic DNA
79	19	1.0	23135	22	AAAS26780 Human immune/haema
80	19	1.0	23135	22	AAK81080 Haemophilus influe
81	19	1.0	1830121	17	AAZ42063 Haemophilus influe
82	19	1.0	1830121	17	AAZ42063 Haemophilus influe

83	18	0.9	18	19	AAV64670	S. pneumoniae PBP2	156	18	0.9	240825	22	AAF24497	Human PG-3 gene.
84	18	0.9	18	19	AAV64671	S. pneumoniae PBP2	c 157	17	0.9	27	19	AAV27438	Streptococcus pneu
c 85	18	0.9	21	22	AAH01135	S. pneumoniae pbp1	158	17	0.9	27	20	AAV25035	Cyclin-dependent k
86	18	0.9	24	22	AAH01204	S. pneumoniae pbp1	159	17	0.9	27	20	AAV25029	Cyclin-dependent k
c 87	18	0.9	383	22	AAI191836	Human polynucleoti	c 160	17	0.9	35	21	AAZ98275	Mouse IP3 sponge g
88	18	0.9	439	22	AAI63993	Human polynucleoti	c 161	17	0.9	36	21	AAZ98274	Mouse IP3 sponge g
89	18	0.9	439	22	AAI63993	Human polynucleoti	c 162	17	0.9	99	21	AAI16155	Human secreted pro
90	18	0.9	461	22	AAI16155	Genomic sequence #	163	17	0.9	130	18	AAV77571	Staphylococcus aur
c 91	18	0.9	491	23	AAI16155	Drosophila melanog	c 164	17	0.9	226	21	AAI16155	Human secreted exp
92	18	0.9	512	23	AAI16155	DNA encoding novel	c 165	17	0.9	234	21	AAI16155	Human secreted exp
93	18	0.9	526	21	AAI16155	cDNA #1069 encodin	c 166	17	0.9	306	21	AAI16155	Neisseria meningit
94	18	0.9	541	22	AAI16155	Arabidopsis thalia	c 167	17	0.9	332	21	AAI16155	Human polynucleoti
95	18	0.9	541	22	AAI16155	Human foetal liver	c 168	17	0.9	340	21	AAI16155	Human secreted pro
96	18	0.9	541	22	AAI16155	Probe #8173 for ge	c 169	17	0.9	341	21	AAI16155	Neisseria meningit
97	18	0.9	541	22	AAI16155	Human brain expres	c 170	17	0.9	402	21	AAI16155	Neisseria meningit
98	18	0.9	541	22	AAI16155	Human bone marrow	c 171	17	0.9	417	22	AAI16155	Human polynucleoti
99	18	0.9	541	22	AAI16155	Probe #7374 for ge	c 172	17	0.9	462	22	AAI16155	Human polynucleoti
c 100	18	0.9	541	22	AAI16155	Probe #11033 used	c 173	17	0.9	500	23	AAI16155	DNA encoding novel
c 101	18	0.9	612	20	AAI16155	Prostate cancer as	c 174	17	0.9	529	21	AAI16155	3' end of cDNA seq
c 102	18	0.9	700	21	AAI16155	Human inflammatory	c 175	17	0.9	560	22	AAI16155	Streptococcus miti
c 103	18	0.9	724	21	AAI16155	Human ORF ORF1628	c 176	17	0.9	560	22	AAI16155	Streptococcus miti
c 104	18	0.9	755	21	AAI16155	Arabidopsis thalia	c 177	17	0.9	561	22	AAI16155	Human digestive sy
c 105	18	0.9	921	21	AAI16155	Arabidopsis thalia	c 178	17	0.9	574	23	AAI16155	Drosophila melanog
c 106	18	0.9	932	20	AAI16155	Arabidopsis cyclin	c 179	17	0.9	636	21	AAI16155	Human ATP binding
c 107	18	0.9	1078	21	AAI16155	Arabidopsis thalia	c 180	17	0.9	639	21	AAI16155	Fusarium venenatum
c 108	18	0.9	1104	23	AAI16155	DNA encoding novel	c 181	17	0.9	706	22	AAI16155	Mammalian vestibul
c 109	18	0.9	1197	21	AAI16155	Arabidopsis thalia	c 182	17	0.9	783	11	AAI16155	Fowl Pox Virus pro
c 110	18	0.9	1200	22	AAI16155	Enterococcus faeca	c 183	17	0.9	794	22	AAI16155	Human neuroblastom
c 111	18	0.9	1229	22	AAI16155	Human colon cancer	c 184	17	0.9	813	22	AAI16155	Human cDNA clone (
c 112	18	0.9	1281	21	AAI16155	Arabidopsis thalia	c 185	17	0.9	834	19	AAI16155	3' cDNA sequence o
c 113	18	0.9	1299	23	AAI16155	E. coli DNA for ce	c 186	17	0.9	834	19	AAI16155	3' fragment of pro
c 114	18	0.9	1319	21	AAI16155	Arabidopsis thalia	c 187	17	0.9	834	21	AAI16155	Human immunogenic
c 115	18	0.9	1348	21	AAI16155	Arabidopsis thalia	c 188	17	0.9	834	22	AAI16155	Human prostate cDN
c 116	18	0.9	1429	16	AAI16155	Neurotrophin-6 cDN	c 189	17	0.9	834	22	AAI16155	Human prostate tum
c 117	18	0.9	1564	22	AAI16155	Human purified sec	c 190	17	0.9	834	22	AAI16155	Human prostate-spe
c 118	18	0.9	1592	21	AAI16155	Human ORF ORF2605	c 191	17	0.9	834	22	AAI16155	Human prostate-spe
c 119	18	0.9	1772	22	AAI16155	Human ORF ORF2605	c 192	17	0.9	834	22	AAI16155	Prostate tumour an
c 120	18	0.9	1788	23	AAI16155	Human MIT-like pro	c 193	17	0.9	971	17	AAI16155	Human clone 41 gen
c 121	18	0.9	1800	22	AAI16155	Drosophila melanog	c 194	17	0.9	1042	22	AAI16155	Human digestive sy
c 122	18	0.9	1800	22	AAI16155	Human cDNA sequenc	c 195	17	0.9	1043	22	AAI16155	Human digestive sy
c 123	18	0.9	1800	22	AAI16155	Human polynucleoti	c 196	17	0.9	1084	22	AAI16155	Human digestive sy
c 124	18	0.9	1800	22	AAI16155	DNA methyltransfer	c 197	17	0.9	1084	22	AAI16155	Streptococcus pneu
c 125	18	0.9	1800	22	AAI16155	Human adult brain	c 198	17	0.9	1095	23	AAI16155	DNA encoding novel
c 126	18	0.9	1800	22	AAI16155	Drosophila melanog	c 199	17	0.9	1134	23	AAI16155	Drosophila melanog
c 127	18	0.9	1800	22	AAI16155	Human secreted pro	c 200	17	0.9	1202	19	AAI16155	H. pylori GHPO 143
c 128	18	0.9	1800	22	AAI16155	Human polynucleoti	c 201	17	0.9	1260	22	AAI16155	Escherichia coli p
c 129	18	0.9	1800	22	AAI16155	DNA encoding novel	c 202	17	0.9	1457	22	AAI16155	Human Nck-2 cDNA
c 130	18	0.9	1800	22	AAI16155	BamWV coat protein	c 203	17	0.9	1460	20	AAI16155	Human secreted pro
c 131	18	0.9	1800	22	AAI16155	DNA encoding novel	c 204	17	0.9	1498	22	AAI16155	Nucleotide sequenc
c 132	18	0.9	1800	22	AAI16155	Drosophila melanog	c 205	17	0.9	1531	21	AAI16155	Human secreted pro
c 133	18	0.9	1800	22	AAI16155	DNA encoding novel	c 206	17	0.9	1713	21	AAI16155	Arabidopsis thalia
c 134	18	0.9	1800	22	AAI16155	DNA encoding novel	c 207	17	0.9	1714	21	AAI16155	Arabidopsis thalia
c 135	18	0.9	1800	22	AAI16155	Drosophila melanog	c 208	17	0.9	1719	23	AAI16155	DNA encoding novel
c 136	18	0.9	1800	22	AAI16155	Long terminal repe	c 209	17	0.9	1812	22	AAI16155	Mouse IP3 sponge p
c 137	18	0.9	1800	22	AAI16155	DNA encoding novel	c 210	17	0.9	1813	22	AAI16155	Human cDNA sequenc
c 138	18	0.9	1800	22	AAI16155	DNA encoding novel	c 211	17	0.9	1906	19	AAI16155	CD30 ligand gene u
c 139	18	0.9	1800	22	AAI16155	Drosophila melanog	c 212	17	0.9	1984	23	AAI16155	Drosophila melanog
c 140	18	0.9	1800	22	AAI16155	Human immune/haema	c 213	17	0.9	1985	22	AAI16155	Fragment of KDR si
c 141	18	0.9	1800	22	AAI16155	DNA encoding human	c 214	17	0.9	2013	22	AAI16155	Human colon cancer
c 142	18	0.9	1800	22	AAI16155	Human reproductive	c 215	17	0.9	2258	23	AAI16155	Drosophila melanog
c 143	18	0.9	1800	22	AAI16155	Human pancreatic c	c 216	17	0.9	2267	20	AAI16155	Human KDR signal t
c 144	18	0.9	1800	22	AAI16155	Human digestive sy	c 217	17	0.9	2347	22	AAI16155	Human cDNA sequenc
c 145	18	0.9	1800	22	AAI16155	Human nervous syst	c 218	17	0.9	2421	21	AAI16155	Arabidopsis thalia
c 146	18	0.9	1800	22	AAI16155	Human nervous syst	c 219	17	0.9	2429	23	AAI16155	Arabidopsis thalia
c 147	18	0.9	1800	22	AAI16155	Human immune/haema	c 220	17	0.9	2520	19	AAI16155	Arabidopsis thalia
c 148	18	0.9	1800	22	AAI16155	DNA methyltransfer	c 221	17	0.9	2594	23	AAI16155	MSENO40-2 promote
c 149	18	0.9	1800	22	AAI16155	Human nervous syst	c 222	17	0.9	2672	24	AAI16155	Drosophila melanog
c 150	18	0.9	1800	22	AAI16155	Human cardiovascular	c 223	17	0.9	2710	21	AAI16155	Human prostate spe
c 151	18	0.9	1800	22	AAI16155	Enterococcus faeca	c 224	17	0.9	2733	22	AAI16155	Group B Streptococ
c 152	18	0.9	1800	22	AAI16155	Human immune/haema	c 225	17	0.9	2741	22	AAI16155	Human cDNA sequenc
c 153	18	0.9	1800	22	AAI16155	Human immune/haema	c 226	17	0.9	2741	22	AAI16155	Human cDNA sequenc
c 154	18	0.9	1800	22	AAI16155	Human immune/haema	c 227	17	0.9	3202	23	AAI16155	Drosophila melanog
c 155	18	0.9	1800	22	AAI16155	Human immune/haema	c 228	17	0.9	3435	22	AAI16155	S. epidermidis gen
						Murine LOBO homolo	c 229	17	0.9	3484	19	AAI16155	Arabidopsis nemato
										3629	23	AAI16155	Drosophila melanog

229	17	0.9	3982	23	ABL03364	Drosophila melanog	c 302	16	0.8	26	24	AAD21908	PCR primer, 2767T
230	17	0.9	4032	22	ABL08104	Drosophila melanog	c 303	16	0.8	30	14	AAQ0767	Rat FC5R primer R2
231	17	0.9	4046	22	ABA88881	Escherichia coli p	c 304	16	0.8	51	22	AAI77097	Human silent SNP c
232	17	0.9	4068	22	AAI60291	Human polynucleoti	c 305	16	0.8	51	22	AAH90159	Human interleukin-
233	17	0.9	4103	22	AAI58505	Human polynucleoti	c 306	16	0.8	133	21	AAH31079	Human secreted pro
234	17	0.9	4261	23	ABL24616	Drosophila melanog	307	16	0.8	140	19	AAX10526	Human biallelic po
235	17	0.9	4765	20	AAX13187	Enterococcus faeca	308	16	0.8	169	21	AAC24235	Human secreted pro
236	17	0.9	4793	23	ABL18712	Drosophila melanog	309	16	0.8	174	22	ABA66272	Human foetal liver
237	17	0.9	4794	23	ABL01874	Drosophila melanog	310	16	0.8	174	22	ABA33338	Probe #11804 for g
238	17	0.9	4805	23	ABL24621	Drosophila melanog	311	16	0.8	174	22	AAK14689	Human brain expres
239	17	0.9	4950	23	ABL21378	Drosophila melanog	312	16	0.8	174	22	AAK40433	Human bone marrow
240	17	0.9	5177	22	AAH56333	Rat soluble adenyli	c 313	16	0.8	175	21	AAC24542	Human secreted pro
241	17	0.9	5177	24	AAH77294	Polynucleotide of	c 314	16	0.8	186	22	AAK19233	Human brain expres
242	17	0.9	5675	23	ABL16334	Drosophila melanog	c 315	16	0.8	186	22	AAK45203	Human bone marrow
243	17	0.9	5820	23	ABL03878	Drosophila melanog	c 316	16	0.8	217	21	AAC22600	Human secreted pro
244	17	0.9	6133	24	AAAS61052	Human gene regulat	c 317	16	0.8	217	21	AAK24337	Human gene signatu
245	17	0.9	6287	23	ABL03876	Drosophila melanog	c 318	16	0.8	248	22	AAK59739	Human immune/haema
246	17	0.9	6623	19	AAV19051	Mitochondrial gene	c 319	16	0.8	256	20	AAK06693	Human immune/haema
247	17	0.9	6623	19	AAV19053	Mitochondrial gene	c 320	16	0.8	256	20	AAK06693	Chromosome 8 deriv
248	17	0.9	7148	23	ABL22002	Drosophila melanog	c 321	16	0.8	258	21	AAA61124	SEN virus genome f
249	17	0.9	7191	23	ABL14461	Drosophila melanog	c 322	16	0.8	264	21	AAC20818	Human secreted pro
250	17	0.9	7574	22	AAK74485	Drosophila melanog	c 323	16	0.8	285	14	AAK18897	Human secreted pro
251	17	0.9	7574	22	AAK74613	Human immune/haema	c 324	16	0.8	285	14	AAQ61137	Human brain Expres
252	17	0.9	8414	22	AAK74613	Human immune/haema	c 325	16	0.8	289	22	AAK58095	Human immune/haema
253	17	0.9	9388	23	ABL24620	DNA encoding novel	c 326	16	0.8	292	17	AAK34605	Probe for detectin
254	17	0.9	9848	21	AAZ98272	Drosophila melanog	c 327	16	0.8	300	20	AAZ12794	Human gene express
255	17	0.9	9871	12	AAQ13593	Mouse IP3 receptor	328	16	0.8	306	21	AAC50168	Arabidopsis thalia
256	17	0.9	10006	24	ABL32036	Human immune syste	c 329	16	0.8	306	21	AAQ10484	HIV Tat_Cys22_Sf16
257	17	0.9	10317	23	ABL14460	Human immune/haema	c 330	16	0.8	310	22	AAK83981	Human immune/haema
258	17	0.9	12714	23	ABL14976	Drosophila melanog	c 331	16	0.8	327	21	AAC45039	Arabidopsis thalia
259	17	0.9	12893	21	AAAB1733	N. meningitidis pa	c 332	16	0.8	332	22	AAK63065	Human immune/haema
260	17	0.9	13982	22	AAK65456	Human immune/haema	c 333	16	0.8	345	22	AAK58749	Human immune/haema
261	17	0.9	13982	22	AAK77762	Human immune/haema	c 334	16	0.8	348	20	AAV86174	Human immune/haema
262	17	0.9	13982	22	AAK84022	Human immune/haema	c 335	16	0.8	349	14	AAQ61138	Human brain Expres
263	17	0.9	16287	24	ABL32672	Human immune syste	c 336	16	0.8	353	15	AAQ44781	EST clone J1168. H
264	17	0.9	17131	21	AAZ60888	DNA encoding a hum	c 337	16	0.8	357	23	AAAS2011	Drosophila anti-ba
265	17	0.9	18920	23	ABL02306	Drosophila melanog	c 338	16	0.8	359	18	AAK75248	Staphylococcus aur
266	17	0.9	21126	20	AAK12960	Enterococcus faeca	c 339	16	0.8	363	15	AAQ44782	Nucleotide sequenc
267	17	0.9	22680	22	AAK66308	Human immune/haema	c 340	16	0.8	363	21	AAC26501	Drosophila anti-ba
268	17	0.9	22680	22	AAK73334	Human immune/haema	c 341	16	0.8	363	21	AAK11914	Human secreted pro
269	17	0.9	22680	22	AAK73344	Human immune/haema	c 342	16	0.8	363	22	AAK59926	Human breast cance
270	17	0.9	22680	22	AAK73625	Human immune/haema	c 343	16	0.8	364	22	AAK20804	Human immune/haema
271	17	0.9	22680	22	AAK73847	Human immune/haema	c 344	16	0.8	380	22	AAK22792	Human breast cance
272	17	0.9	22680	22	AAK73934	Human immune/haema	c 345	16	0.8	388	22	AAK09550	Human breast cance
273	17	0.9	22680	22	AAK78350	Human immune/haema	c 346	16	0.8	389	22	AAK22558	Human breast cance
274	17	0.9	22788	23	ABL20208	Drosophila melanog	c 347	16	0.8	390	22	AAK74773	Human immune/haema
275	17	0.9	24908	22	AAK81665	Human immune/haema	c 348	16	0.8	392	21	AAC07428	Human secreted pro
276	17	0.9	29973	22	AAK29814	Human cytoskeletal	c 349	16	0.8	393	22	ABA11994	Human nervous syst
277	17	0.9	31034	23	ABL13678	Drosophila melanog	c 350	16	0.8	400	18	AAV76468	Staphylococcus aur
278	17	0.9	47475	21	AAAB1465	N. meningitidis pa	c 351	16	0.8	402	22	AAK22302	Human immunoglobul
279	17	0.9	48203	22	AAK70161	Human immune/haema	c 352	16	0.8	406	22	AAH29374	Drosophila melanog
280	17	0.9	48203	22	AAK81663	Human immune/haema	c 353	16	0.8	417	21	AAA61158	SEN virus genome f
281	17	0.9	48203	22	AAK82628	Human immune/haema	c 354	16	0.8	421	22	AAK13689	Human breast cance
282	17	0.9	48204	22	AAK70164	Human immune/haema	c 355	16	0.8	423	20	AAK98755	EST clone HK669.
283	17	0.9	48204	22	AAK81666	Human immune/haema	c 356	16	0.8	423	21	AAK16425	Human prostate can
284	17	0.9	48204	22	AAK82630	Human immune/haema	c 357	16	0.8	432	22	AAK60625	Human immune/haema
285	17	0.9	49431	23	ABL13742	Drosophila melanog	c 358	16	0.8	432	22	AAC26669	Human secreted pro
286	17	0.9	49767	21	AAA81458	N. meningitidis pa	c 359	16	0.8	435	22	AAK76539	Human immune/haema
287	17	0.9	54863	22	AAK86025	Human immune/haema	c 360	16	0.8	438	24	ABK16439	Gram positive bact
288	17	0.9	54877	22	AAK86026	Human immune/haema	c 361	16	0.8	438	24	ABK16464	Gram positive bact
289	17	0.9	58407	19	AAV21210	Methanococcus jann	c 362	16	0.8	449	22	ABA46193	Human breast cell
290	17	0.9	111309	20	AAK20250	Borrelia burgdorfe	c 363	16	0.8	449	22	ABA56733	Human foetal liver
291	17	0.9	151826	21	AAK22291	BAC containing rep	c 364	16	0.8	449	22	ABA36359	Probe #4825 for ge
292	17	0.9	172325	21	AAK21613	Neisseria meningit	c 365	16	0.8	449	22	AAK04865	Human brain expres
293	17	0.9	349980	21	AAK21608	Neisseria meningit	c 366	16	0.8	449	22	AAK30393	Human bone marrow
294	17	0.9	349980	21	AAK21612	Neisseria meningit	c 367	16	0.8	449	22	AAI15001	Probe #4934 for ge
295	17	0.9	349980	22	AAH41226	Pyrococcus abyssi	c 368	16	0.8	449	22	AAI36346	Probe #5032 used t
296	17	0.9	837096	21	AAA81489	N. meningitidis pa	c 369	16	0.8	449	22	AAI04766	Probe #4757 used t
297	17	0.9	1038602	20	AAZ01425	Complete genome se	c 370	16	0.8	451	21	AAC24241	Human secreted pro
298	17	0.9	1664976	19	AAV21209	Methanococcus jann	c 371	16	0.8	456	23	AAAS4920	Staphylococcus aur
299	16	0.8	20	22	AAK26694	Human Nck-2 phosph	c 372	16	0.8	456	23	AAK05318	Staphylococcus aur
300	16	0.8	21	22	AAH26607	Human Mda-7 gene p	c 373	16	0.8	458	11	AAQ05318	Sequence encoding
301	16	0.8	24	24	ABK16600	CXCR2 receptor PCR	c 374	16	0.8	465	22	AAQ11644	Sequence encoding
												AAK06466	Human brain expres

c 375	16	0.8	465	22	AAK32147	Human bone marrow	448	16	0.8	700	22	AAH92519	Human inflammatory
376	16	0.8	470	22	ABA53690	Human foetal liver	c 449	16	0.8	700	22	AAH93259	Human inflammatory
377	16	0.8	470	22	ABA23443	Probe #1909 for ge	450	16	0.8	707	24	AAH92861	Human G protein-co
378	16	0.8	470	22	AAK01954	Human brain expres	c 451	16	0.8	714	22	AAF25573	H. pylori HPC005 e
379	16	0.8	470	22	AAK27412	Human bone marrow	c 452	16	0.8	722	22	AAH06929	Human CDNA clone (
c 380	16	0.8	471	22	ABA43971	Human breast cell	c 453	16	0.8	724	21	AAF13210	Aspergillus oryzae
c 381	16	0.8	471	22	ABA54424	Human foetal liver	454	16	0.8	732	22	AAI95781	Human neuroblastom
c 382	16	0.8	471	22	ABA24206	Probe #2672 for ge	455	16	0.8	745	21	AAZ97441	Human prostate can
c 383	16	0.8	471	22	AAH08088	Human breast cancer	c 456	16	0.8	747	22	AAH35622	Human musculoskele
c 384	16	0.8	473	21	AAC37200	Arabidopsis thalia	c 457	16	0.8	748	22	AAI95530	Human neuroblastom
385	16	0.8	475	22	ABA59103	Human foetal liver	c 458	16	0.8	766	19	AAI98565	DNA encoding a S.
386	16	0.8	475	22	AAK07282	Human brain expres	c 459	16	0.8	767	22	AAH22556	Human cDNA encodin
387	16	0.8	475	22	AAK33042	Human bone marrow	c 460	16	0.8	773	19	AAV61971	Human mu-opioid re
c 388	16	0.8	475	22	AAI38845	Probe #7531 used t	c 461	16	0.8	773	19	AAV61972	Human mu-opioid re
c 389	16	0.8	481	22	AAH01692	Human reproductive	c 462	16	0.8	773	19	AAV61973	Human mu-opioid re
c 390	16	0.8	481	22	AAH26763	Human genomic DNA	c 463	16	0.8	773	19	AAV61974	Human mu-opioid re
c 391	16	0.8	481	22	AAH26764	Human genomic DNA	c 464	16	0.8	773	19	AAV61975	Human mu-opioid re
c 392	16	0.8	486	20	AAH84692	Human metastatic m	c 465	16	0.8	773	19	AAV61976	Human mu-opioid re
c 393	16	0.8	495	22	AAH18567	Streptococcus pneu	c 466	16	0.8	773	19	AAV61977	Human mu-opioid re
c 394	16	0.8	495	22	AAH26264	Human cDNA encodin	c 467	16	0.8	773	19	AAV61978	Human mu-opioid re
c 395	16	0.8	505	22	ABA19800	Human nervous syst	c 468	16	0.8	786	20	AAH98717	Human validated ca
c 396	16	0.8	505	22	ABA19801	Human nervous syst	c 469	16	0.8	791	22	AAH08302	Human CDNA clone (
397	16	0.8	510	22	AAH92639	Human inflammatory	c 470	16	0.8	799	21	AAH50278	Arabidopsis thalia
c 398	16	0.8	517	22	AAH12278	Human cytochrome p	c 471	16	0.8	801	22	AAH91789	Human cDNA 5'-end
399	16	0.8	526	21	AAH45585	Arabidopsis thalia	c 472	16	0.8	810	22	AAH02070	Streptococcus pneu
c 400	16	0.8	528	22	AAH68562	Human immune/haema	c 473	16	0.8	816	22	AAH93709	Human cDNA clone r
401	16	0.8	528	22	AAH68563	Human immune/haema	c 474	16	0.8	816	22	AAH06754	Human cDNA clone (
c 402	16	0.8	534	22	AAH14944	Human breast cancer	c 475	16	0.8	826	23	AAH64722	DNA encoding novel
c 403	16	0.8	534	22	AAH84812	Human immune/haema	c 476	16	0.8	831	21	AAH05573	Streptococcus pneu
c 404	16	0.8	539	21	AAH97996	Human colon cancer	c 477	16	0.8	831	22	AAH90732	CFE 36 coding sequ
c 405	16	0.8	541	21	AAH41938	Arabidopsis thalia	c 478	16	0.8	834	22	AAH81674	Human immune/haema
c 406	16	0.8	549	21	AAH12602	Aspergillus oryzae	c 479	16	0.8	834	22	AAH00086	Ehrlichia canis nu
407	16	0.8	549	21	AAH94795	Cat flea hindgut p	c 480	16	0.8	849	21	AAH15233	Trichoderma reesei
c 408	16	0.8	550	22	AAH12277	Human cytochrome p	c 481	16	0.8	854	22	AAH90845	2CFE 36 coding seq
c 409	16	0.8	570	21	AAH50166	Arabidopsis thalia	c 482	16	0.8	860	21	AAH45493	Arabidopsis thalia
410	16	0.8	574	19	AAH60062	Nucleic acid GII5c	c 483	16	0.8	865	21	AAH38605	Arabidopsis thalia
c 411	16	0.8	578	22	AAH16635	Human nervous syst	c 484	16	0.8	868	22	AAH03333	Human cDNA clone (
c 412	16	0.8	578	22	AAH02068	Human cardiovascular	c 485	16	0.8	875	18	AAH84032	DNA encoding a Sta
c 413	16	0.8	579	22	AAH02068	Streptococcus pneu	c 486	16	0.8	875	19	AAH53425	DNA encoding a Sta
c 414	16	0.8	586	21	AAH08837	Fusarium venenatum	c 487	16	0.8	880	21	AAC39609	Arabidopsis thalia
c 415	16	0.8	590	21	AAC33873	Arabidopsis thalia	c 488	16	0.8	881	21	AAC60864	Mouse immortalizat
c 416	16	0.8	592	22	AAH23079	Osteoarthritis tis	c 489	16	0.8	890	19	AAH60003	Nucleic acid G256
c 417	16	0.8	594	22	AAH35599	Human foetal liver	c 490	16	0.8	891	21	AAH50477	Arabidopsis thalia
c 418	16	0.8	594	22	AAH12125	Human brain expres	c 491	16	0.8	891	21	AAH34964	Arabidopsis thalia
c 419	16	0.8	594	22	AAH37842	Human bone marrow	c 492	16	0.8	904	21	AAH34964	Arabidopsis thalia
c 420	16	0.8	594	22	AAH43715	Probe #12401 used	c 493	16	0.8	930	23	AAH56115	Salmonella typhi D
c 421	16	0.8	595	19	AAH41690	Nucleotide sequenc	c 494	16	0.8	951	22	AAH31990	Human olfactory re
c 422	16	0.8	601	21	AAC40754	Arabidopsis thalia	c 495	16	0.8	959	20	AAH24950	Melanoma kinase DN
c 423	16	0.8	601	22	AAH28979	Drosophila melanog	c 496	16	0.8	960	23	AAH91367	DNA encoding novel
c 424	16	0.8	607	22	AAH46903	Human G protein-co	c 497	16	0.8	978	21	AAH52211	Human UGT2B15 exon
c 425	16	0.8	609	21	AAC77030	Human ORFX ORF2585	c 498	16	0.8	985	24	AAH34320	Human immune syste
c 426	16	0.8	609	21	AAC46165	Arabidopsis thalia	c 499	16	0.8	988	20	AAH20183	Enterococcus faeca
c 427	16	0.8	611	22	AAH23800	Human breast cancer	c 500	16	0.8	1029	23	AAH53309	Haemophilus influe
c 428	16	0.8	618	22	AAH11628	Human foetal liver	c 501	16	0.8	1041	22	AAH81429	Escherichia coli p
c 429	16	0.8	618	22	AAH19966	Human brain expres	c 502	16	0.8	1041	23	AAH52472	E. coli DNA for ce
c 430	16	0.8	618	22	AAH46003	Human bone marrow	c 503	16	0.8	1054	20	AAH84944	Human secreted pro
c 431	16	0.8	618	22	AAH151918	Probe #26064 used	c 504	16	0.8	1055	21	AAH44244	Arabidopsis thalia
c 432	16	0.8	622	20	AAH87425	EST clone BR559.	c 505	16	0.8	1055	22	AAH71308	Human immune/haema
c 433	16	0.8	625	21	AAH08607	Fusarium venenatum	c 506	16	0.8	1059	21	AAC46105	Arabidopsis thalia
c 434	16	0.8	643	19	AAH98650	DNA encoding a S.	c 507	16	0.8	1062	21	AAC33736	Arabidopsis thalia
c 435	16	0.8	645	22	AAH26765	Human genomic DNA	c 508	16	0.8	1078	22	AAH57854	Human immune/haema
c 436	16	0.8	646	22	AAH26767	Human genomic DNA	c 509	16	0.8	1092	20	AAH20182	Enterococcus faeca
c 437	16	0.8	654	22	AAH34120	Human cDNA encodin	c 510	16	0.8	1112	18	AAH84063	DNA encoding a S.
c 438	16	0.8	673	21	AAH13747	Aspergillus oryzae	c 511	16	0.8	1146	21	AAH05814	Group B Streptococ
c 439	16	0.8	673	23	AAH83171	DNA encoding novel	c 512	16	0.8	1155	22	AAC39187	Arabidopsis thalia
c 440	16	0.8	676	21	AAC47722	Arabidopsis thalia	c 513	16	0.8	1161	22	AAH63759	Drosophila gustato
c 441	16	0.8	679	22	ABA11164	Human nervous syst	c 514	16	0.8	1185	21	AAC43410	Arabidopsis thalia
c 442	16	0.8	682	22	AAH03483	Human cDNA clone (c 515	16	0.8	1186	13	AAH29184	RubisCo promoted h
c 443	16	0.8	683	22	AAH21664	Human breast cancer	c 516	16	0.8	1194	22	AAH52232	Human AFP protein
c 444	16	0.8	697	21	AAH13471	Aspergillus oryzae	c 517	16	0.8	1232	17	AAH34614	Probe for detectin
c 445	16	0.8	699	22	AAH92104	Human inflammatory	c 518	16	0.8	1257	23	AAH78461	DNA encoding novel
c 446	16	0.8	700	22	AAH92105	Human inflammatory	c 519	16	0.8	1262	21	AAH26404	Human secreted pro
c 447	16	0.8	700	22	AAH92518	Human inflammatory	c 520	16	0.8	1279	22	AAH58681	Human polynucleoti
										1287	22	ABA08873	Human secreted pro

c 521	16	0.8	1295	19	AAV11659	Rat spleen ICAM-4	c 594	16	0.8	1721	11	AAQ06319	Synthetic gene enc
c 522	16	0.8	1295	22	AAV19330	Rat spleen partial	c 595	16	0.8	1721	12	AAQ13164	Gene encoding t-PA
c 523	16	0.8	1295	19	AAV19330	Partial rat ICAM-R	c 596	16	0.8	1721	12	AAQ13165	Gene encoding t-PA
c 524	16	0.8	1314	18	AAV63436	Mangosteen Class I	c 597	16	0.8	1721	20	AAH07566	Homo sapiens fetal
c 525	16	0.8	1314	18	AAV63436	Mangosteen Class I	c 598	16	0.8	1722	20	AAH07566	C glutamicum codin
c 526	16	0.8	1314	20	AAV71847	Mangosteen C18:1 a	c 599	16	0.8	1727	14	AAQ46262	Encodes modified t
c 527	16	0.8	1317	22	AAV29058	cdNA encoding for	c 600	16	0.8	1732	23	AAH77812	DNA encoding novel
c 528	16	0.8	1353	17	AAV41706	Lymphocyte specifi	c 601	16	0.8	1738	22	AAH35761	Human musculoskele
c 529	16	0.8	1353	21	AAV39862	Plasmodium ovale p	c 602	16	0.8	1739	22	AAH35760	Human musculoskele
c 530	16	0.8	1387	22	AAV29132	cdNA encoding for	c 603	16	0.8	1743	20	AAV72028	Adenovirus wild-ty
c 531	16	0.8	1400	15	AAQ70142	Human cartilage li	c 604	16	0.8	1743	21	AAH37052	Human PRO1315 (UNQ
c 532	16	0.8	1420	22	AAV06495	Arabidopsis thalia	c 605	16	0.8	1743	22	AAH92095	DNA encoding prote
c 533	16	0.8	1435	19	AAV16466	Nucleotide sequenc	c 606	16	0.8	1743	22	AAH54274	DNA encoding prote
c 534	16	0.8	1449	23	ABL26303	Drosophila melanog	c 607	16	0.8	1746	21	AAH01223	Adenovirus 5 (Ad5)
c 535	16	0.8	1456	24	AAV594703	Rat secreted facto	c 608	16	0.8	1746	22	AAH31329	Recombinant adenov
c 536	16	0.8	1472	19	AAV34660	Rodent cdNA clone	c 609	16	0.8	1748	16	AAH099008	Interleukin 8 rece
c 537	16	0.8	1472	19	AAV11662	Rat ICAM-4 clone E	c 610	16	0.8	1749	22	AAH17031	Human cdNA sequenc
c 538	16	0.8	1472	19	AAV19333	Rat ICAM-4 partial	c 611	16	0.8	1750	21	AAH20991	Human low adenosin
c 539	16	0.8	1472	20	AAV36492	Rat ICAM coding se	c 612	16	0.8	1750	21	AAH34869	Human adenosine re
c 540	16	0.8	1472	20	AAV08979	Rat ICAM-4 coding	c 613	16	0.8	1751	21	AAH27133	Human inflammation
c 541	16	0.8	1474	21	AAV39630	Arabidopsis thalia	c 614	16	0.8	1751	14	AAQ34490	Insert in pcgpl75
c 542	16	0.8	1477	22	AAH58737	Human polynucleoti	c 615	16	0.8	1759	11	AAQ01277	DNA fragment carry
c 543	16	0.8	1479	23	ABL19975	Drosophila melanog	c 616	16	0.8	1801	18	AAH67080	Mouse oligodendroc
c 544	16	0.8	1481	13	AAQ23496	Encodes pre t-PA v	c 617	16	0.8	1801	22	AAH89708	Mouse oligodendroc
c 545	16	0.8	1486	22	AAH65954	Nucleotide sequenc	c 618	16	0.8	1801	24	AAH99895	Mouse ischaemic co
c 546	16	0.8	1497	22	AAH60467	Human polynucleoti	c 619	16	0.8	1807	21	AAH20992	Human low adenosin
c 547	16	0.8	1502	22	AAH79031	Human immune/haema	c 620	16	0.8	1807	21	AAH34870	Human adenosine re
c 548	16	0.8	1502	22	AAH79033	Human immune/haema	c 621	16	0.8	1812	14	AAQ34489	Human adenosine re
c 549	16	0.8	1510	14	AAQ38747	Human IL-8 recepto	c 622	16	0.8	1812	14	AAQ49417	Insert in pcgpl76
c 550	16	0.8	1510	21	AAH20993	Human low adenosin	c 623	16	0.8	1818	21	AAH93819	Cytochrome P450 ho
c 551	16	0.8	1510	21	AAH34871	Human adenosine re	c 624	16	0.8	1818	21	AAH93821	Modified fibre pro
c 552	16	0.8	1512	23	ABL22001	Drosophila melanog	c 625	16	0.8	1824	14	AAH47842	Modified fibre pro
c 553	16	0.8	1515	19	AAV52022	Helicobacter poly	c 626	16	0.8	1837	21	AAH34806	Arabidopsis thalia
c 554	16	0.8	1524	22	AAH88103	Human FLEHT-34 nu	c 627	16	0.8	1848	21	AAH93820	Modified fibre pro
c 555	16	0.8	1528	13	AAQ28715	Wild-type AOX2 pro	c 628	16	0.8	1848	21	AAH93822	Modified fibre pro
c 556	16	0.8	1528	13	AAQ28717	Mutant AOX2 promot	c 629	16	0.8	1884	21	AAH15985	Human prostate can
c 557	16	0.8	1528	15	AAQ63285	Natural AOX2 promo	c 630	16	0.8	1886	22	AAH09268	Human IL-8R B homo
c 558	16	0.8	1528	15	AAQ63289	Mutant AOX2 promot	c 631	16	0.8	1893	21	AAH93823	Modified fibre pro
c 559	16	0.8	1528	16	AAH081992	Alcohol oxidase (A	c 632	16	0.8	1893	21	AAH93824	Modified fibre pro
c 560	16	0.8	1529	22	AAH160388	Human polynucleoti	c 633	16	0.8	1920	23	AAH92476	DNA encoding novel
c 561	16	0.8	1530	22	AAH52352	S. epidermidis ope	c 634	16	0.8	1923	21	AAH93825	Modified fibre pro
c 562	16	0.8	1536	21	AAH43428	Arabidopsis thalia	c 635	16	0.8	1923	21	AAH93826	Modified fibre pro
c 563	16	0.8	1538	21	AAQ47897	Arabidopsis thalia	c 636	16	0.8	1926	21	AAH61125	Modified fibre pro
c 564	16	0.8	1547	13	AAQ28718	Mutant AOX2 promot	c 637	16	0.8	1931	19	AAH03142	SEN virus genome f
c 565	16	0.8	1547	15	AAQ63291	Mutant AOX2 promot	c 638	16	0.8	1931	20	AAH27413	Aspergillus nidula
c 566	16	0.8	1553	22	AAH58602	Human polynucleoti	c 639	16	0.8	1944	23	AAH55607	A. nidulans phytas
c 567	16	0.8	1565	21	AAH38033	Arabidopsis thalia	c 640	16	0.8	1963	22	AAH08826	Streptococcus pneu
c 568	16	0.8	1568	19	AAH68326	Human STIF cdNA	c 641	16	0.8	1973	24	AAH26494	Human class II cyt
c 569	16	0.8	1579	23	AAH86135	DNA encoding novel	c 642	16	0.8	1976	21	AAH95206	Sheep alpha (1, 3)
c 570	16	0.8	1588	21	AAH29800	Sequence upstream	c 643	16	0.8	1976	21	AAH95206	Human UDP-glucuron
c 571	16	0.8	1595	21	AAH22417	Human secreted pro	c 644	16	0.8	1983	22	AAH18571	Human cdNA sequenc
c 572	16	0.8	1605	11	AAQ06317	Synthetic gene enc	c 645	16	0.8	2009	22	AAH05410	Human secreted pro
c 573	16	0.8	1605	13	AAQ23498	Encodes t-PA varia	c 646	16	0.8	2018	21	AAH26865	Essential Staphylo
c 574	16	0.8	1605	14	AAQ46260	Synthetic tPA gene	c 647	16	0.8	2018	22	AAH08016	Staphylococcus aur
c 575	16	0.8	1609	10	AAH90742	Chemically synthe	c 648	16	0.8	2018	22	AAH91547	Staphylococcus aur
c 576	16	0.8	1616	10	AAH90742	Synthetic gene cod	c 649	16	0.8	2024	21	AAH94052	Haemophilus contort
c 577	16	0.8	1620	22	AAH29145	Polynucleotide seq	c 650	16	0.8	2039	22	AAH45110	cdNA encoding nove
c 578	16	0.8	1639	16	AAH74059	Human interleukin-	c 651	16	0.8	2066	23	AAH81269	DNA encoding novel
c 579	16	0.8	1639	20	AAH14999	DNA encoding an in	c 652	16	0.8	2076	22	AAH16467	Nucleotide sequenc
c 580	16	0.8	1639	22	AAH02920	Human shear stress	c 653	16	0.8	2076	22	AAH56326	DNA encoding Chlam
c 581	16	0.8	1652	21	AAH79011	Human secreted pro	c 654	16	0.8	2079	21	AAH61156	SEN virus genome f
c 582	16	0.8	1660	21	AAH75849	Human ORFX ORF1404	c 655	16	0.8	2088	17	AAH36895	Candida albicans l
c 583	16	0.8	1660	22	AAH08323	Human Cdc42-bindin	c 656	16	0.8	2088	22	AAH07947	Human amyloid prec
c 584	16	0.8	1673	24	AAH562329	cdNA sequence #116	c 657	16	0.8	2091	20	AAH18839	Streptococcus equi
c 585	16	0.8	1686	22	AAH57487	Human liver cell s	c 658	16	0.8	2106	22	AAH94726	Human full-length
c 586	16	0.8	1689	23	AAH52961	DNA encoding novel	c 659	16	0.8	2106	22	AAH94726	Human full-length
c 587	16	0.8	1700	19	AAH32893	Human interleukin-	c 660	16	0.8	2114	23	AAH67980	DNA encoding novel
c 588	16	0.8	1700	20	AAH10787	Human tumour supre	c 661	16	0.8	2136	22	AAH60006	Human synthetase #
c 589	16	0.8	1705	22	AAH46121	Human DNA encoding	c 662	16	0.8	2149	22	AAH99560	Human protein enco
c 590	16	0.8	1705	22	AAH44264	Human PRO3301 nucl	c 663	16	0.8	2149	21	AAH99560	Human protein enco
c 591	16	0.8	1706	21	AAH40491	Human fetal kidney	c 664	16	0.8	2161	21	AAH21788	Human breast and o
c 592	16	0.8	1718	21	AAH90052	Melanoma different	c 665	16	0.8	2176	23	ABL10759	Drosophila melanog
c 593	16	0.8	1719	21	AAH51095	Arabidopsis thalia	c 666	16	0.8	2184	23	ABL15025	Drosophila melanog
										2195	23	ABL16866	Drosophila melanog

667	16	0.8	2197	22	ABA07287	Human pancreatic c	740	16	0.8	3373	23	ABL06164	Drosophila melanog
668	16	0.8	2197	22	AAK89932	Human digestive sy	c 741	16	0.8	3394	22	AAH54265	S. epidermidis gen
669	16	0.8	2214	18	AAH84275	Rat ICAM-4 cDNA RT	c 742	16	0.8	3416	22	AAH81784	Human differential
670	16	0.8	2214	19	AAV34658	RT-PCR clone for c	743	16	0.8	3428	22	AAI07433	Human reproductive
671	16	0.8	2214	19	AAV11660	Rat ICAM-4 cDNA fr	744	16	0.8	3476	23	ABL162729	Human breast or ov
672	16	0.8	2214	19	AAV19331	Rat ICAM-4 RT-PCR	c 745	16	0.8	3476	23	ABL19970	Drosophila melanog
673	16	0.8	2214	20	AAK36497	Rat ICAM coding se	c 746	16	0.8	3483	23	ABL25574	Drosophila melanog
674	16	0.8	2214	20	AAV08977	Rat ICAM-4 coding	747	16	0.8	3516	19	AAV57515	zcytor7 cytokine r
675	16	0.8	2247	18	AAV25134	H. pylori inner me	748	16	0.8	3516	22	AAH07643	Human cDNA encodin
676	16	0.8	2247	18	AAV24920	H. pylori ORF 05a	749	16	0.8	3516	22	AAH22815	Human IL-20 recept
677	16	0.8	2260	22	AAH40861	DNA encoding human	750	16	0.8	3532	21	AAH49919	Mouse homologue of
678	16	0.8	2260	22	AAH06549	Human reproductive	751	16	0.8	3562	22	AAH137385	Human musculoskele
679	16	0.8	2268	21	AAK38342	DNA encoding human	752	16	0.8	3573	23	ABL12105	Drosophila melanog
680	16	0.8	2269	21	AAK39419	Human ets2 coding	753	16	0.8	3595	23	ABL26302	Drosophila melanog
681	16	0.8	2286	22	AAH26595	Human melanoma dif	c 754	16	0.8	3598	22	AAF93803	Human cDNA encodin
682	16	0.8	2296	23	AAH77847	DNA encoding novel	c 755	16	0.8	3601	22	AAH54151	S. epidermidis gen
683	16	0.8	2314	21	AAK76745	Human ORFX ORF2300	756	16	0.8	3608	21	AAH99916	cDNA encoding huma
684	16	0.8	2329	19	AAK14596	H. pylori GHP0 175	757	16	0.8	3608	22	AAH46141	Human DNA encoding
685	16	0.8	2360	22	AAH07734	Human secreted pro	758	16	0.8	3635	17	AAH06761	Human c-mer protoo
686	16	0.8	2368	21	AAH26851	Essential Staphylo	759	16	0.8	3635	23	ABL19872	Drosophila melanog
687	16	0.8	2368	22	AAH08002	Staphylococcus aur	760	16	0.8	3661	23	ABL22000	Drosophila melanog
688	16	0.8	2368	22	AAH91533	Staphylococcus aur	761	16	0.8	3692	22	AAH33758	Human colon cancer
689	16	0.8	2373	20	AAH13400	Enterococcus faeca	c 762	16	0.8	3746	21	AAH99512	DNA encoding a mai
690	16	0.8	2396	23	ABL25394	Drosophila melanog	c 763	16	0.8	3746	21	AAH99527	DNA encoding a mai
691	16	0.8	2400	23	AAH51712	Staphylococcus aur	c 764	16	0.8	3749	23	ABL19974	Drosophila melanog
692	16	0.8	2403	17	AAH12562	S.aureus topoisome	c 765	16	0.8	3773	21	AAH299494	DNA encoding a mai
693	16	0.8	2403	17	AAH12573	S.aureus mutant gr	c 766	16	0.8	3799	24	AAH16458	Corn cDNA encoding
694	16	0.8	2403	23	AAH01330	Staphylococcus aur	c 767	16	0.8	3801	22	AAH05465	Human reproductive
695	16	0.8	2403	23	AAH54404	Staphylococcus aur	c 768	16	0.8	3866	22	AAH54616	S. epidermidis gen
696	16	0.8	2403	23	AAH55401	Staphylococcus aur	c 769	16	0.8	3881	21	AAH61269	Human secreted pro
697	16	0.8	2405	22	AAH02351	Human RNA metaboli	c 770	16	0.8	3891	21	AAH76424	Human ORFX ORF1979
698	16	0.8	2469	21	AAH65740	Streptococcus pneu	c 771	16	0.8	3935	21	AAH98913	Human pancreatic c
699	16	0.8	2469	21	AAH05811	Group B Streptococ	772	16	0.8	3955	23	ABL25064	Drosophila melanog
700	16	0.8	2469	22	AAH00038	Streptococcus agal	773	16	0.8	3957	23	AAH78289	DNA encoding novel
701	16	0.8	2472	21	AAH65741	Streptococcus pneu	774	16	0.8	4080	23	ABL03766	Drosophila melanog
702	16	0.8	2478	22	AAH00036	Streptococcus pyog	c 775	16	0.8	4099	23	AAH86847	DNA encoding novel
703	16	0.8	2480	22	AAH15977	Human cDNA sequenc	c 776	16	0.8	4145	24	ABH99868	Mouse ischaemic co
704	16	0.8	2550	19	AAH34661	Human polynucleoti	c 777	16	0.8	4223	22	AAH58965	Human polynucleoti
705	16	0.8	2550	19	AAH11663	SEN virus genome f	c 778	16	0.8	4223	22	AAH160751	Human polynucleoti
706	16	0.8	2550	19	AAH19334	Rat ICAM-4 clone 7	779	16	0.8	4248	21	AAH56910	S. aureus_RUSA266
707	16	0.8	2550	20	AAH08980	Rat brain ICAM-4 p	c 780	16	0.8	4254	23	ABL13025	Drosophila melanog
708	16	0.8	2567	21	AAH35376	Rat ICAM-4 coding	c 781	16	0.8	4301	23	ABL12688	Drosophila melanog
709	16	0.8	2568	22	AAH15184	Absorptive hyperca	c 782	16	0.8	4305	20	AAH87945	Candida cloacae fa
710	16	0.8	2579	22	AAH60794	Human cDNA sequenc	c 783	16	0.8	4373	18	AAH72719	Pyruvate:flavodoxi
711	16	0.8	2646	21	AAH61122	Human polynucleoti	c 784	16	0.8	4394	22	AAH35762	Human musculoskele
712	16	0.8	2665	23	AAH74027	SEN virus genome f	c 785	16	0.8	4400	22	AAH35763	Human musculoskele
713	16	0.8	2760	21	AAH51152	DNA encoding novel	c 786	16	0.8	4468	22	AAH58364	Human polynucleoti
714	16	0.8	2760	22	AAH07710	Arabidopsis thalia	c 787	16	0.8	4495	17	AAH32143	Helicobacter-speci
715	16	0.8	2780	22	AAH61239	Human secreted pro	c 788	16	0.8	4505	16	AAH88164	Tomato TGTR1 ethy
716	16	0.8	2784	22	AAH18856	Soybean raffinose	c 789	16	0.8	4566	19	AAH59048	Tomato ethylene re
717	16	0.8	2785	23	AAH77796	Human nervous syst	c 790	16	0.8	4590	22	AAH18208	Human cDNA sequenc
718	16	0.8	2854	22	AAH18607	DNA encoding novel	c 791	16	0.8	4600	23	ABL22914	Drosophila melanog
719	16	0.8	2885	22	AAH159008	Human nervous syst	c 792	16	0.8	4601	21	AAH97542	Streptomyces albul
720	16	0.8	2885	22	AAH159008	Human polynucleoti	c 793	16	0.8	4641	23	ABL16126	Drosophila melanog
721	16	0.8	2936	22	AAH21832	Genomic DNA encodi	c 794	16	0.8	4652	22	AAH60150	Human polynucleoti
722	16	0.8	2988	18	AAH84245	Human collagen gen	c 795	16	0.8	4705	23	ABL10758	Drosophila melanog
723	16	0.8	2988	19	AAH34657	Rat ICAM-4 cDNA.	c 796	16	0.8	4718	23	ABL21466	Drosophila melanog
724	16	0.8	2988	19	AAH11655	Rat intercellular	c 797	16	0.8	4771	23	ABL22038	Drosophila melanog
725	16	0.8	2988	19	AAH19326	Rat ICAM-4 cDNA.	c 798	16	0.8	4783	23	ABL09156	Drosophila melanog
726	16	0.8	2988	20	AAH36486	Rat ICAM-4 coding	c 799	16	0.8	4800	19	AAH16558	Rat ICAM-4 genomic
727	16	0.8	2988	20	AAH08991	Rat ICAM-4 coding	c 800	16	0.8	4900	19	AAH19329	Rat ICAM overlappi
728	16	0.8	2997	23	AAH59339	DNA encoding novel	c 801	16	0.8	4900	20	AAH36488	Rat ICAM-1 coding
729	16	0.8	3016	22	AAH54928	S. epidermidis gen	c 802	16	0.8	4900	22	AAH91946	Partial rat ICAM-R
730	16	0.8	3025	18	AAH74489	Staphylococcus aur	c 803	16	0.8	4938	23	ABL04108	Drosophila melanog
731	16	0.8	3062	22	AAH54645	S. epidermidis gen	c 804	16	0.8	5077	19	AAH34659	Rat cDNA clone for
732	16	0.8	3118	23	ABL04190	Drosophila melanog	c 805	16	0.8	5077	19	AAH16661	Rat ICAM-4 clone E
733	16	0.8	3121	22	AAH54245	S. epidermidis gen	c 806	16	0.8	5077	19	AAH19332	Rat ICAM-4 partial
734	16	0.8	3204	22	AAH94369	Human full-length	c 807	16	0.8	5077	20	AAH36491	Rat ICAM coding se
735	16	0.8	3223	23	AAH593540	DNA encoding novel	c 808	16	0.8	5077	20	AAH08978	Rat ICAM-4 coding
736	16	0.8	3252	23	ABL18460	Drosophila melanog	c 809	16	0.8	5163	24	ABL33248	Human immune syste
737	16	0.8	3324	22	AAH16336	Human cDNA sequenc	c 810	16	0.8	5176	19	AAH04017	Human multiple mve
738	16	0.8	3324	22	AAH16336	Human cDNA sequenc	c 811	16	0.8	5215	20	AAH91105	Group B Streptococ
739	16	0.8	3347	21	AAH61121	SEN virus genome f	c 812	16	0.8	5217	21	AAH256330	CKS-P66-CKS fusion

C 813	16	0.8	5347	24	AAS94844	Human DNA sequence
C 814	16	0.8	5470	23	AAQ38596	Human glucose regu
C 815	16	0.8	5823	14	ABL12104	Drosophila melanog
C 816	16	0.8	5839	22	AAV46335	Hydra head activat
C 817	16	0.8	5976	22	AAS46301	Tumour suppressor
C 818	16	0.8	5983	23	AAS74036	DNA encoding novel
C 819	16	0.8	5987	24	ABL33563	Human immune syste
C 820	16	0.8	5997	22	AAI66084	Leishmania major h
C 821	16	0.8	6016	22	AAH57390	Human skeletal mus
C 822	16	0.8	6096	22	AAS26766	Human genomic DNA
C 823	16	0.8	6171	19	AAV52170	Streptococcus pneu
C 824	16	0.8	6398	23	ABL25726	Drosophila melanog
C 825	16	0.8	6508	23	ABL09474	Drosophila melanog
C 826	16	0.8	6618	22	ABA18101	Human nervous syst
C 827	16	0.8	6666	12	AAQ10416	Mutant protease ge
C 828	16	0.8	6675	12	AAQ10414	Mutant protease ge
C 829	16	0.8	6675	12	AAQ10415	Mutant protease ge
C 830	16	0.8	6675	12	AAQ10411	Mutant protease ge
C 831	16	0.8	6675	12	AAQ10412	Mutant protease ge
C 832	16	0.8	6675	12	AAQ10413	Mutant protease ge
C 833	16	0.8	6675	12	AAQ10417	Mutant protease ge
C 834	16	0.8	6693	12	AAQ10871	Mutant protease ge
C 835	16	0.8	6711	12	AAQ10870	Mutant protease ge
C 836	16	0.8	6806	18	AAV74783	Staphylococcus aur
C 837	16	0.8	6882	22	ABA14861	Human nervous syst
C 838	16	0.8	6882	22	ABA15569	Human nervous syst
C 839	16	0.8	6979	23	AAS81429	DNA encoding novel
C 840	16	0.8	6990	22	AAI58278	Human polynucleoti
C 841	16	0.8	7027	22	AAI60064	Human LXR1 longer
C 842	16	0.8	7081	18	AAT74201	Staphylococcus aur
C 843	16	0.8	7104	23	AAS51998	Staphylococcus aur
C 844	16	0.8	7107	19	AAV32371	Complete sequence
C 845	16	0.8	7148	21	AAAS9042	Nucleotide sequenc
C 846	16	0.8	7148	24	ABA94251	Nucleotide sequenc
C 847	16	0.8	7156	20	AAI12966	Enterococcus faeca
C 848	16	0.8	7337	20	AAI13079	Enterococcus faeca
C 849	16	0.8	7337	20	AAI13079	DNA encoding novel
C 850	16	0.8	7389	22	ABL37088	Drosophila melanog
C 851	16	0.8	7417	23	ABL13024	Staphylococcus aur
C 852	16	0.8	7434	23	AAS52179	Staphylococcus aur
C 853	16	0.8	7437	23	AAS55232	Staphylococcus aur
C 854	16	0.8	7449	24	ABL32276	Human immune syste
C 855	16	0.8	7469	19	AAV32372	Complete sequence
C 856	16	0.8	7469	21	AAAS9043	Nucleotide sequenc
C 857	16	0.8	7469	24	ABA94252	Nucleotide sequenc
C 858	16	0.8	7567	23	ABL12481	Drosophila melanog
C 859	16	0.8	7584	24	AAS95251	Long terminal repe
C 860	16	0.8	7726	22	AAI66083	Leishmania donovan
C 861	16	0.8	7737	23	ABL02204	Drosophila melanog
C 862	16	0.8	7865	23	ABL06402	Drosophila melanog
C 863	16	0.8	7902	23	ABL06398	Drosophila melanog
C 864	16	0.8	7958	21	AAZ34835	Mouse integrin sub
C 865	16	0.8	7960	21	AAAS9072	Nucleotide sequenc
C 866	16	0.8	7960	24	ABA94274	Nucleotide sequenc
C 867	16	0.8	7989	21	AAAS9075	Nucleotide sequenc
C 868	16	0.8	7989	24	ABA94277	Nucleotide sequenc
C 869	16	0.8	8155	18	AAV74374	Staphylococcus aur
C 870	16	0.8	8365	20	AAZ20056	Plasmodium falcipa
C 871	16	0.8	8370	22	AAS46713	Tumour suppressor
C 872	16	0.8	8383	21	AAAS9071	Nucleotide sequenc
C 873	16	0.8	8383	24	ABA94273	Nucleotide sequenc
C 874	16	0.8	8471	22	AAS32667	Human genomic DNA
C 875	16	0.8	8480	20	AAZ27521	Interleukin-2 rece
C 876	16	0.8	8484	21	AAAS9091	Nucleotide sequenc
C 877	16	0.8	8748	23	ABL12592	Drosophila melanog
C 878	16	0.8	8821	23	ABL10984	Drosophila melanog
C 879	16	0.8	8907	22	ABA21312	Human nervous syst
C 880	16	0.8	8907	22	ABA21313	Human nervous syst
C 881	16	0.8	8907	22	AAI03164	Human reproductive
C 882	16	0.8	8907	22	AAI03165	Human reproductive
C 883	16	0.8	8930	19	AAV22834	Haemophilus paraga
C 884	16	0.8	8962	24	ABL2686	Human immune syste
C 885	16	0.8	9001	22	AAK84675	Human immune/haema
C 886	16	0.8	9120	22	ABA14655	Human nervous syst
C 887	16	0.8	9120	22	ABA14656	Human nervous syst
C 888	16	0.8	9180	22	AAI04447	Human reproductive
C 889	16	0.8	9206	24	ABL33675	Human immune syste
C 890	16	0.8	9319	21	AAK20995	Human low adenosin
C 891	16	0.8	9319	21	AAK34873	Human adenosine re
C 892	16	0.8	9345	23	ABL09624	Drosophila melanog
C 893	16	0.8	9741	24	ABL33322	Human immune syste
C 894	16	0.8	9789	17	AAI14852	cDNA encoding plas
C 895	16	0.8	10133	24	ABL32458	Human immune syste
C 896	16	0.8	10218	22	AAK77457	Human immune/haema
C 897	16	0.8	10296	23	ABL15024	Drosophila melanog
C 898	16	0.8	10610	19	AAV32375	Complete sequence
C 899	16	0.8	10610	21	AAAS9051	Nucleotide sequenc
C 900	16	0.8	10610	24	ABA94260	Nucleotide sequenc
C 901	16	0.8	10694	23	ABL07242	Drosophila melanog
C 902	16	0.8	10776	23	ABL33804	Drosophila melanog
C 903	16	0.8	10835	22	AAK79845	Human immune/haema
C 904	16	0.8	10982	22	AAS46264	DNA encoding novel
C 905	16	0.8	11034	24	AAD26491	Sheep alpha (1, 3)
C 906	16	0.8	11050	18	AAV74407	Staphylococcus aur
C 907	16	0.8	11176	22	AAS01080	Mouse FOSB DNA. M
C 908	16	0.8	11176	24	AAD22528	Mouse FOSB DNA. M
C 909	16	0.8	11532	22	ABA07434	Human pancreatic c
C 910	16	0.8	11532	22	AAK91165	Human digestive sy
C 911	16	0.8	12103	22	AAI05045	Human reproductive
C 912	16	0.8	12103	22	AAS33419	DNA encoding human
C 913	16	0.8	12225	18	AAT89257	Human bg gene (sho
C 914	16	0.8	12278	22	ABA14585	Human nervous syst
C 915	16	0.8	12278	22	ABA14664	Human nervous syst
C 916	16	0.8	12283	12	AAQ10353	Nucleotide sequenc
C 917	16	0.8	12343	22	AAS46240	DNA encoding novel
C 918	16	0.8	12460	24	ABL32069	Human immune syste
C 919	16	0.8	12616	18	AAT89256	Human bg gene (lon
C 920	16	0.8	12789	24	AAS16876	Human interleukin
C 921	16	0.8	13003	23	ABL12480	Drosophila melanog
C 922	16	0.8	13605	22	AAI37081	Human musculoskele
C 923	16	0.8	13767	23	ABL15058	Drosophila melanog
C 924	16	0.8	14333	22	AAK79846	Human immune/haema
C 925	16	0.8	14455	19	AAV32374	Complete sequence
C 926	16	0.8	14455	21	AAAS9050	Nucleotide sequenc
C 927	16	0.8	14455	24	ABA94259	Nucleotide sequenc
C 928	16	0.8	14458	23	ABL10220	Drosophila melanog
C 929	16	0.8	14537	22	AAS46355	Tumour suppressor
C 930	16	0.8	14657	23	ABL18606	Drosophila melanog
C 931	16	0.8	14775	24	ABI99535	Mouse ischaemic co
C 932	16	0.8	15644	22	AAS34611	Human DNA for a no
C 933	16	0.8	15870	20	AAI13035	Enterococcus faeca
C 934	16	0.8	16891	20	AAK37084	MEFV gene sequence
C 935	16	0.8	17674	24	ABL33344	Human immune syste
C 936	16	0.8	18177	10	AAN90490	DNA of human retin
C 937	16	0.8	18303	20	AAK04502	Human retinoblasto
C 938	16	0.8	18475	19	AAV52171	Streptococcus pneu
C 939	16	0.8	18598	24	ABL32386	Human immune syste
C 940	16	0.8	18869	23	ABL04818	Drosophila melanog
C 941	16	0.8	20752	22	AAK75098	Human immune/haema
C 942	16	0.8	21591	20	AAI13047	Enterococcus faeca
C 943	16	0.8	21732	22	AAS39820	Genomic sequence #
C 944	16	0.8	21732	22	AAK30176	Human digestive sy
C 945	16	0.8	21799	22	ABA07343	Human pancreatic c
C 946	16	0.8	21799	22	AAS32770	Human immune/haema
C 947	16	0.8	22401	22	AAK71911	Human immune/haema
C 948	16	0.8	22402	22	AAK84676	Human immune/haema
C 949	16	0.8	23054	22	AAK84676	Genomic sequence #
C 950	16	0.8	23452	22	AAS42122	Genomic sequence #
C 951	16	0.8	23457	22	AAS42121	Genomic sequence #
C 952	16	0.8	23458	22	AAS42120	Tumour suppressor
C 953	16	0.8	24259	22	AAS46691	Human musculoskele
C 954	16	0.8	24699	22	AAI136054	Human nervous syst
C 955	16	0.8	25003	22	ABA19679	Human immune/haema
C 956	16	0.8	25131	22	AAK84544	Human immune/haema
C 957	16	0.8	26370	23	ABL07082	Drosophila melanog
C 958	16	0.8	26879	23	AAS59524	Propionibacterium

QY	541	ATCGAATGAAAATCTAAGGCTACTCTCTGCTGACAGTGTGAGAAAGCAAGTCAATAC	600
Dp	541	atctggaatatgaaaataatcaagctctacatctctctgtgaacagtaatgagaagcagttcaatac	600
QY	601	ACCAATTACGATGAGGACTATCAAAAGTCTCAAAATCAGCAACATTAATTACCGCTTCATATGA	660
Dp	601	accaattactgataggactacaagaagttccaatcagaagaatatacccttgcttaacatgga	660
QY	661	TAAATTACCTCAAGGAAGTCAATCAATCAAGTTGAGAGAAGCAAGGCTTTAACTTACTAC	720
Dp	661	taattactctaaggaagtcatacatcaagttgaagaagaacagcgctataactactcac	720
QY	721	AACGGGATGGATGTCTACACAAATGTACAAACGAAGAGCTCAAAAACATCTGTGGGATAT	780
Dp	721	aacgggataggatgctctacacaaatgtagaaccaagaagctcaaaaacatctgtyggatat	780
QY	781	TTACAATACAGACGAATACGTTGCTCTATCCACAGATGATTTGCAAGTGGCTTACCAT	840
Dp	781	ttacaatacagaagataagcttgccctacacagatgatatgcaagtcgctctaccat	840
QY	841	TGTTGATGTTTCTAACGGTAAAGTCATTGCCACGCTAGSAGCAGCCATCACTCAAGTAA	900
Dp	841	tgttgatgcttctaacagtgtaaagtcattgccacgctagaagacgcatacagtaagtaa	900
QY	901	TGTTTCCTTGGAATTAACCAACGACTATCAAAACAAACCCGACGCTGGGATCAACATTA	960
Dp	901	tgtttccttggaattaaaccaacgactatcaaaaacaaacccgacgctgggatacaatatgaa	960
QY	961	ACCGATCACAGACTATGCTCTCTGCTTGGAGTACGGGTGTTCACGATTCAGCTACTAT	1020
Dp	961	accgatcacagactatgctctcctgcttggagtagtgctcagttcaactgctactat	1020
QY	1021	CGTTACAGATGAGACCCCTATACATACCCCTGGGACAAATACCTCGTTTAACTGGGATAG	1080
Dp	1021	cgttacagatgagacccctatatcataccctgggacaataactcctgttataacttggagtag	1080
QY	1081	GGGCTACTTTGGAGACATCACTCCCTTGCAATACGCCCTGCAACATGCGCAAAAGTCCAGC	1140
Dp	1081	gggctactttggagaacatacatcacttgaataagcccttgacaacaatcggaagctccagc	1140
QY	1141	CGTGGAAACTCTAAACAAGGTGGGACTCAACCGCGCCAAAGACTTTTCTTAATGCTTAGG	1200
Dp	1141	cgtggaaactcttaacaagaagtggaactcaacgcgcgaagactttcttaaatgcttagg	1200
QY	1201	AATGACTATCCCAAGATTCTACTACTCAATGCCAATTTCAATTAACACACACCGATCAGA	1260
Dp	1201	aatgactatacccaagattctactctaccataatgcatltaagtaacacacacgaatcaga	1260
QY	1261	CAAAAAATATGAGCAGATAGTAAAGAATGGCTGCTGCTTAGCGCTTGCCATATGG	1320
Dp	1261	caaaaaatatgagcagaatagtgaaaagaatggctgctgcttagcgcttgccaatag	1320
QY	1321	TGGAACTTACTATTAACCAATGTATTCATTAAGACGCTTTAGTATGGAGAGTAAAA	1380
Dp	1321	tggaaacttactataaaccaaatgtatatccaataaagtcgtctttagtgatggagtagaaa	1380
QY	1381	AGAGTCTCTAATGTGCGAATCTGTGCCATGAGAGGAAGACAGCGCTATATGATGACGA	1440
Dp	1381	agaagtctctaatgtcgaactcgctgcaatgaagaaacagcactatagtagcga	1440
QY	1441	CATGATGAAGACGTCTGTAGCTAATGGAACGTGAGCAAAATGCGTATCTTGCTTGGCTCC	1500
Dp	1441	catgatgaagacgtctgtagctaatggaaactgagcaaaaatgctatctgtgtgctcc	1500
QY	1501	TCAGGCTGTGTAAACAGGAACCTCTTAACATATACAGACGAGAAATGTAAAAACCAATCAA	1560
Dp	1501	tcagagctgtgtaaaacaggaacctctaactatacagacgaggaatgtgaaaaccaatcaa	1560
QY	1561	GACCTCTCAATTTGTAGCACCTGTATGAACATATTTGCTGGCTATACGGTAAATATTCAT	1620
Dp	1561	gacctctcaatttgtagcacctgtatgaacatatttgcgtgctatacgttaatatccaat	1620

QY	1621	GGCTGTATGAGCAGAGCTATTCTTAACCGTCTGACACACACTTGTAGGCAATGGCTTACGGT	1680
Db	1621	ggctgtatlgagcagagctattctttaaaccgtctgcacacacttglaggaacaagggcttaccggt	1680
QY	1681	CGCGGCCAAATTTTACCGGCTGTATGATGACCTACCTGTCGTGAAGGAAGCAATCCGAAGA	1740
Db	1681	cgcgcgccaagttaaccgctctatgatgacctaccgctctgaaggaagaaaccagaaga	1740
QY	1741	TTTGGATATATACAGAGGGGGCTCTACAGAAATGGAGAATTCGTATTAAAATATGTCGTCG	1800
Db	1741	tttggatatataccagagggtctctacagaaatlgagaattcgtattttaaataatgtygctcg	1800
QY	1801	TTTCTACTGTGGAACCTCACCTGCTCCACACACACCCCATATACCTGAATGAATTCACGCTATC	1860
Db	1801	tttctactgtggaactcacacctctctccacaacaaccccatcaactgtaaatgttcaagctc	1860
QY	1861	ATTCAGATGTTCACACTTCACAGTCTTACCTCAACCACTCCAGACAAATTAATAGTACAGC	1920
Db	1861	atctcagatgttcaacttccacagctctagcttcaaccactccaagacaataataatagtagac	1920
QY	1921	TACCAATCTCAACATTAATATAGCAACACATCAAAATACACCCCTGTATCAACAAATCAGAA	1980
Db	1921	taccaatctctaaacaataatagcaacaacaataatatacaacccctgtatcaacaataatcagaa	1980
QY	1981	TTCTTCACACGACACACCA	1999
Db	1981	ttcttcaaccagcacaca	1999

RESULT 2
 AAV52278/c
 ID AAV52278 standard; DNA; 10711 BP.
 XX
 AC AAV52278;
 XX
 DT 23-OCT-1998 (first entry)
 XX
 DE Streptococcus pneumoniae genome fragment SEQ ID NO:145.
 XX
 KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
 KW computer readable medium; vaccine; pharmaceutical composition; ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO9818931-A2.
 XX
 PD 07-MAY-1998.
 XX
 PF 30-OCT-1997; 97WO-US19588.
 XX
 PR 31-OCT-1996; 96US-0029960.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
 PI Kunsch CA, Rosen CA;
 XX
 DR WPI; 1998-272225/24.
 XX
 PT Computer-readable medium with recorded Streptococcus pneumoniae
 PT polynucleotide sequences - useful in diagnostic kits and assays, and
 PT pharmaceutical compositions and vaccines for Streptococcus
 PT pneumoniae
 XX
 PS Claim 1; Page 972-978; 14099p; English.
 XX
 CC The present invention describes a computer readable medium which has
 CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
 CC recorded on it, or a representative fragment or a sequence at least 95%
 CC identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
 CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
 CC Streptococcus pneumoniae. The present invention also describes an
 CC isolated nucleic acid molecule encoding a homologue of any of the

CC fragments of the *S. pneumoniae* genome (SEQ ID NO:1 to 391) where the
CC nucleic acid library is produced by a process comprising: (a) screening
CC a genomic DNA library using as a probe a target sequence defined by any
CC of the sequences in SEQ ID NO:1 to 391, identifying members of the
CC library which contain sequences that hybridize to the target sequence and
CC isolating the nucleic acid molecules from the members; or (b) isolating
CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
CC molecules whose nucleotide sequence is homologous to amplification
CC primers derived from the fragment of the *S. pneumoniae* genome to prime
CC the amplification and isolating the amplified sequences; The computer
CC readable medium can be used in a computer-based system for identifying
CC fragments of the *S. pneumoniae* genome of commercial importance, or
CC expression modulating fragments of the *S. pneumoniae* genome. Products
CC from the present invention can be used in diagnosis kits and assays, and
CC pharmaceutical compositions and vaccines for *S. pneumoniae*.

XX Sequence 10711 BP; 3280 A; 2056 C; 2412 G; 2963 T; 0 other:

Query Match 100.0%; Score 1999; DB 19; Length 10711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATTAATCACTACTGCTGACTGGGTTCTGACGCCGCTCAA 60
DB 9767 TAAATCTACGACAAATTAATCACTACTGCTGACTGGGTTCTGACGCCGCTCAA 9708
QY 61 TGCCCAAGTAATGATTTCCACAGATTTGGTTAAGSCAATGTTCTATCGAAGCA 120
DB 9707 TGCCCAAGTAATGATTTCCACAGATTTGGTTAAGSCAATGTTCTATCGAAGCA 9648
QY 121 TCGCTTCTCGACACAGGGGGATTTGATACATCCGTATCCTGGAGCTTCTTGCCAA 180
DB 9647 TCGCTTCTCGACACAGGGGGATTTGATACATCCGTATCCTGGAGCTTCTTGCCAA 9588
QY 181 TCTGCAAGCAATTCCTCCAGGTGATCACTCTCACCACAGATTGATTAAGTTGAC 240
DB 9587 TCTGCAAGCAATTCCTCCAGGTGATCACTCTCACCACAGATTGATTAAGTTGAC 9528
QY 241 TTTACTTTCACTTCGACTCCGACGACATTTCTCGTAAGGCTCGAGAGGTTGGTT 300
DB 9527 TTTACTTTCACTTCGACTCCGACGACATTTCTCGTAAGGCTCGAGAGGTTGGTT 9468
QY 301 AGCGATTCACTTGAACAAAAGCAACCAAGCAAGAAATTTGACTATATATAATA 360
DB 9467 AGCGATTCACTTGAACAAAAGCAACCAAGCAAGAAATTTGACTATATATAATA 9408
QY 361 GGTCTACATGCTTAATGGGAATGGAATGACAGACAGCTCAAACTATATGTTAA 420
DB 9407 GGTCTACATGCTTAATGGGAATGGAATGACAGACAGCTCAAACTATATGTTAA 9348
QY 421 AGACCTCAATTAATTTAGTTTACCTAGTTAGCTTCTGGTGGAGTGGCTGAGCAC 480
DB 9347 AGACCTCAATTAATTTAGTTTACCTAGTTAGCTTCTGGTGGAGTGGCTGAGCAC 9288
QY 481 AAACCAATATGACCCCTATTTCATCCAGAGACGCCAAGCCGCAAACTTGCTTT 540
DB 9287 AAACCAATATGACCCCTATTTCATCCAGAGACGCCAAGCCGCAAACTTGCTTT 9228
QY 541 ATCTGAAATGAAATTAAGGCTACATCTGCTGAACAGTATGAGAAGCAGTCAATAC 600
DB 9227 ATCTGAAATGAAATTAAGGCTACATCTGCTGAACAGTATGAGAAGCAGTCAATAC 9168
QY 601 ACCAATTAATGATGATCAAGTCAATATGAGCAAGTAATTAACCTGCTTACATGGA 660
DB 9167 ACCAATTAATGATGATCAAGTCAATATGAGCAAGTAATTAACCTGCTTACATGGA 9108
QY 661 TAAATTAATGAGGAGTCAATCAATCAAGTTGAAGAGAAACAGGCTATTAACCTACTCAC 720
DB 9107 TAAATTAATGAGGAGTCAATCAATCAAGTTGAAGAGAAACAGGCTATTAACCTACTCAC 9048
QY 721 AACTGGAGTATGCTTACCAAAATGTGACCAAGAGCTCAAAACATCTGTGGATAT 780
DB 721 AACTGGAGTATGCTTACCAAAATGTGACCAAGAGCTCAAAACATCTGTGGATAT 780

DB 9047 AACTGGAGTATGCTTACCAAAATGTGACCAAGAGCTCAAAACATCTGTGGATAT 8988
QY 761 TTAATAATGAGCAATACGTTGCTTATCCAGAGATGAATGCAAGCTGCTTACAT 840
DB 8987 TTAATAATGAGCAATACGTTGCTTATCCAGAGATGAATGCAAGCTGCTTACAT 8928
QY 841 TGTGATGTTTCTTAACGTAAGTCAATTTGCCACCTAGAGACAGCCCATGATCAAGTAA 900
DB 8927 TGTGATGTTTCTTAACGTAAGTCAATTTGCCACCTAGAGACAGCCCATGATCAAGTAA 8868
QY 901 TGTTCCTTTGGGAATTAACCAAGCATGTAAGAAACAACCCGAGCTGGGATCACTATGAA 960
DB 8867 TGTTCCTTTGGGAATTAACCAAGCATGTAAGAAACAACCCGAGCTGGGATCACTATGAA 8808
QY 961 ACCGATCAAGACATATGCTCTGCTGGATGAGTGTGATGATTAACCTGCTACTAT 1020
DB 8807 ACCGATCAAGACATATGCTCTGCTGGATGAGTGTGATGATTAACCTGCTACTAT 8748
QY 1021 CGTTACAGATGAGCCCTATTAACCTCCCTGGGACAAATACTCTGTTTATACTGGATAG 1080
DB 8747 CGTTACAGATGAGCCCTATTAACCTCCCTGGGACAAATACTCTGTTTATACTGGATAG 8688
QY 1081 GGGCTACTTTGGCAACATCACTTGGCAATACGCCCTGCAACAATCGGGAACGCTCCAGC 1140
DB 8687 GGGCTACTTTGGCAACATCACTTGGCAATACGCCCTGCAACAATCGGGAACGCTCCAGC 8628
QY 1141 CGTGAACACTTAACCAAGGTCGAGCTCAACCGGCCCAAGACTTCTCTTAATGTTCTAGG 1200
DB 8627 CGTGAACACTTAACCAAGGTCGAGCTCAACCGGCCCAAGACTTCTCTTAATGTTCTAGG 8568
QY 1201 AATGACTACCCAAAGTATTCATCTCAATATGCCATTTCAAGTAACAAACCAAGTACA 1260
DB 8567 AATGACTACCCAAAGTATTCATCTCAATATGCCATTTCAAGTAACAAACCAAGTACA 8508
QY 1261 CAAAAAATATGAGCAAGTATGTAAGAAATGCGTGTGCTTACGCTCTTGGAAATGG 1320
DB 8507 CAAAAAATATGAGCAAGTATGTAAGAAATGCGTGTGCTTACGCTCTTGGAAATGG 8448
QY 1321 TGGAACTTACTATTAACCAATGTAATTCATCAATTAAGTGTGTTAGTGGAGTGAATA 1380
DB 8447 TGGAACTTACTATTAACCAATGTAATTCATCAATTAAGTGTGTTAGTGGAGTGAATA 8388
QY 8447 TGGAACTTACTATTAACCAATGTAATTCATCAATTAAGTGTGTTAGTGGAGTGAATA 8388
QY 1381 AGAGTTCTTAATGTCGGAACCTGCTGCCATGAAGAAAGACAGACCTATATGATGACGA 1440
DB 8387 AGAGTTCTTAATGTCGGAACCTGCTGCCATGAAGAAAGACAGACCTATATGATGACGA 8328
QY 1441 CATGATGAAAAACAGTCTTGAATATGGAATGGAACGAAATGCGCTATTTGCTGGCTCCC 1500
DB 8327 CATGATGAAAAACAGTCTTGAATATGGAATGGAACGAAATGCGCTATTTGCTGGCTCCC 8268
QY 1501 TCAGGCTGTGTAACAGGAACCTCTACTATATACAGAGAGAAATGGAACCAACATCAA 1560
DB 8267 TCAGGCTGTGTAACAGGAACCTCTACTATATACAGAGAGAAATGGAACCAACATCAA 8208
QY 1561 GACCTCTCAATTTGTAGCCACTGATGAATATTTGCTGGCTATACGCGTAATATTTCAAT 1620
DB 8207 GACCTCTCAATTTGTAGCCACTGATGAATATTTGCTGGCTATACGCGTAATATTTCAAT 8148
QY 1621 GCGTGTATGAGACAGGCTATTTCTAACCGCTGACACCACTTGTAGGCAATGGCTTACGGT 1680
DB 8147 GCGTGTATGAGACAGGCTATTTCTAACCGCTGACACCACTTGTAGGCAATGGCTTACGGT 8088
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGAGCTTACCTGTGGAAGGAGCAATCCAGAGA 1740
DB 8087 CGCTGCCAAAGTTTACCGCTCTATGATGAGCTTACCTGTGGAAGGAGCAATCCAGAGA 8028
QY 1741 TTGGAATATACAGAGGCGCTTACAGAAATGGAATTCGATTTAAAAATGGTGTCTCG 1800
DB 8027 TTGGAATATACAGAGGCGCTTACAGAAATGGAATTCGATTTAAAAATGGTGTCTCG 7968
QY 1801 TTCTAGTGTGAACTACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 1860
DB 7967 TTCTAGTGTGAACTACCTGCTCCACAAACCCCATCACTGAAAGTTCAAGCTCATC 7908

QY 1861 ATGAGATAGTTCACATTCAGTCTAGTCAACCTCCAGCACAATATAGTAGCAG 1920
|||||
DB 7907 ATGAGTAGTCACTTCACAGTCTAGTCAACCTCCAGCACAATATAGTAGCAG 7848
QY 1921 TACCAATCTTAACAAATATACGCAACATCAATATACACCCCTGATCAACAAATACGAA 1980
|||||
DB 7847 TACCAATCTTAACAAATATACGCAACATCAATATACACCCCTGATCAACAAATACGAA 7788
QY 1981 TCTCAACGACGACACCA 1999
|||||
DB 7787 TCTCAACGACGACACCA 7769

RESULT 3
AAT08027
ID AAT08027 standard; DNA; 2048 BP.
AC AAT08027;
XX
XX 03-DEC-1996 (first entry)
DE S. pneumoniae penicillin binding protein 1A soluble variant DNA.
XX
XX Penicillin binding protein; PBP 1A; bifunctional protein;
KM transglycosylase; transpeptidase; identification; assay; inhibitor;
KM antibiotic resistant; bacteria; soluble variant; protein structure;
KW X-ray crystallography; determination; ds.
XX
OS Streptococcus pneumoniae.
FH
FH Key Location/Qualifiers
FT CDS 1..2046
FT /tag= a
XX
XX GB2290792..A.
XX
XX 10-JAN-1996.
PD
XX
XX 29-JUN-1995; 95GB-0013306.
PF
XX
XX 24-NOV-1994; 94SE-0004072.
PR 01-JUL-1994; 94IN-0000580.
XX
XX (ASTR) ASTRA AB.
XX
XX Balganesch TS, Town CM;
PI
XX
XX WPI: 1996-042232/05.
DR P-PSDB; AAM04359.
XX
XX Sol. deriv. of bifunctional penicillin binding protein (BPBP) -
PT opt. lack transglycosylase activity, useful to identify and assay
PT for antibodies or cpds. which bind BPBP
XX
XX Claim 11; Pages 65-68; 108bp; English.
XX
XX The present sequence encodes a soluble variant of the S. pneumoniae
CC penicillin binding protein (PBP) 1A, which lacks the 1st 38 amino
CC acid residues of the wild type protein. Wild type PBP is a
CC bifunctional protein, which binds the cell membrane when expressed
CC in a bacterial cell, having transglycosylase and transpeptidase
CC activities. The variant protein (NCIMB 40665) in conjunction with
CC a labelled anti-bifunctional PBP monoclonal antibody, can be used
CC to identify and assay for cpds. which bind bifunctional PBP. Such
CC cpds., as inhibitors of bifunctional PBP have a potential use in
CC therapeutic cpds. which inhibit the growth of antibiotic resistant
CC bacteria. The soluble variant may also be used in X-ray
CC crystallography.
XX
XX Sequence 2048 BP; 667 A; 498 C; 390 G; 493 T; 0 other;

Query Match 71.9%; Score 1438; DB 17; Length 2048;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATTAATAATCACTCATCTGCTGACTTGGTTGTAACGCCGCTCAA 60
|||||
DB 48 taatactcagcaataataataataactcactcgtctgacttggctcgaagcgcgtcaa 107
QY 61 TGGCCAGGTAATGATATTCGCCACAGATTGGTTAAGGCAATGGTTTATGCAACACCA 120
|||||
DB 108 tgcacaagcctaagataltccacagatttggtaaggaactcgttctacagaagacca 167
QY 121 TCGCTTCTTCGACACAGGGGATGATACATCCGTATCCCTGGAGCTTTCTTGGCAA 180
|||||
DB 168 tgcctcttcgaccacaggggattgataccatccgtatcccttgaggcttcttgcgaa 227
QY 181 TCTGCAAGCAATTCCTCCAGAGGTGATCACTGCACCCAAACAGTTGATTAAATTGAC 240
|||||
DB 228 tctgcaagcaatctccctccaaagtgtatcagctcactcaacagttgaltgaattgac 287
QY 241 TTACTTTCACTTGGACTTCCGACCACTATTTCTCGTAAGGTCAGCAAGCTTGTT 300
|||||
DB 288 ttaacttcaactcgtactccgaccagactattctcgtlaaggtccaagagcttgglt 347
QY 301 AGCGATTCAAGTTAGAACAAACCAACCAAGCAAGAAATCTTACCCTACTATTAATAA 360
|||||
DB 348 agcgattcagttagaacaaacaaacaaacaaacaaagaaatcttgcactactataataa 407
QY 361 GGTCTACATGTCTTAATGGCACTATGGAATGACAGACAGCTCAAAACTATATGTTAA 420
|||||
DB 408 ggtctacatgtcttaatggcaactatggaaatgacagacagctcaaaactatgtaa 467
QY 421 AGACCTCAATTAATTTAAGTTTACCTACGTAGGCTTGCGGCTGGAATGCTCAGGACAC 480
|||||
DB 468 agacctcaatataatagttaccctacgttagtccgttcggtcgtgaatgctcgaagcacc 527
QY 481 AAACCAATATGACCCCTATTTCACATCCAGAGGACGCCAAGAGCGGAACTTGGTCTT 540
|||||
DB 528 aaaccaatattgaccctatttcaatccagaggaagccaaagccgcgaagaacttgglt 587
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAACAGCTCAATAC 600
|||||
DB 588 atctgaatgaaaaatcaaggctacacatctcgtcgaacagttagaaagcaactacatc 647
QY 601 ACCAATTACGATGAGCTCAAAAGTCTCAAAATAGCAAGTAATTAACCTCTTACATGGA 660
|||||
DB 648 accaattacgtagggctacaagttccaatcaatcagaagtaattacatccgtctacatgga 707
QY 661 TAATTACCTCAAGGAATCATCATCAAGTTGAAGAAGAAACAGGCTATTAACCTACTAC 720
|||||
DB 708 taattacctcaaggaaagtcataatcaatcaagttgaagaagaacagctataactactc 767
QY 721 AACTGGATGAGTGTATACAAATGTAGACCAAGAGCTCAAAACATCTGTGGATAT 780
|||||
DB 768 aactggatgagtgatctacacaatgttagaccaagaagctcaaaaacatcttgggatat 827
QY 781 TTACATACAGACGATATAGTGGCTATCCAGAGCATGTAATTCAGAGTGGCTCTACCAT 840
|||||
DB 828 ttacaatacagagaataatagttgcttaccagaagctagatctgcaagcttctacat 887
QY 841 TGTGTAGTGTCTTAACAGGTAAAGTCATTTGCCCGCTAGAGAGCAGCCATCAGTCAAGTAA 900
|||||
DB 888 tgttgttcttcaaggtaaagtcattgcccagctagaagacgcatacagctcaagtaa 947
QY 901 TGTTCCTTCGGAATTAACCAAGCACTAGAAACAAACCGGATGGGATCAACTATGAA 960
|||||
DB 948 tgttcttcctcgaaatcaacaaagctagaacaaacccgacttgggttcaactatgaa 1007
QY 961 ACCGATCAGAGACTATGCTGCTGCTGGAGTACGGTGTCTACGATTCAACTCTACTAT 1020
|||||
DB 1008 accgatcaagactatgctgctgctggagtacgggtgtctacgattcaactctactat 1067
QY 1021 CGTTACAGATGAGCCCTATTAACCTCCGTGGACAAATACTCCTGTTTATTAACGTGGATAG 1080

Db 1068 cgtcacgagatgagccataactaccctgggacaataaccctgttataactgagtag 1127
 QY 1081 GGGCTACTTTGGCAACATCATCTGCAATACGCCCTGCAACATCGGAAACGCCAGC 1140
 Db 1128 gggctacttggcaacatacactctgcaataagccctgcaacaacacgcgaacgcccagc 1187
 QY 1141 CGTGAACACTTAAACAAGGTGGAGCTCAACCGCCGCAAGACTTCTTAAATGGTCTAGG 1200
 Db 1188 cgtggaaccttaacaagaagtcggagctcaacccgcgaagacttctctaaatggtccgg 1247
 QY 1201 AATGACTTACCCAGTATTCTACTACTCAATGCCATTTCAGATACACACCGAATCAGA 1260
 Db 1248 aatgactaccgaagtattactactcaaatgaccttcaagtaacaacacgaatcaga 1307
 QY 1261 CAAAAAATPATGAGCAGTGTGTAAGAGTGGCTGCTAGCGCTGCTTGGCAATGG 1320
 Db 1308 caaaaataatggaacagatgagaaagatgagctgctgcttaacgctgaccttgaatgg 1367
 QY 1321 TGAACCTTACTATTAACCAATGTATATCATTAAGTGTCTTTAGTATGGAGTGAATA 1380
 Db 1368 tggagacttactataacacatgatatccataaagtcgtctttagtgatggagtgaaaa 1427
 QY 1381 AGAGTTCTTAATGTGGAACTGCTGCCATGAGAGAAAGCAGACGCTATATGATGACGA 1440
 Db 1428 agagttcttaattgctggaactcgtccatgaagaacgacgactatgatgacgga 1487
 QY 1441 CATGATTAACCAATCTTGTGAACTGAGCAAGAAATCCATATGCTTGGCTGCC 1500
 Db 1488 catgatatgaacagctctgagcttatgagaaagatgagctatctgcttggctccc 1547
 QY 1501 TCAGGCTGTGTAACAGGAACCTTAACTATACAGAGAGAAATGTAACACCATCA 1560
 Db 1548 tcagctgctgaataacaggaaccttaactatacagacgaggaatgtaaaacacatcaa 1607
 QY 1561 GACCTCTCAATTTGTACACCTGATGAACATTTTGGCTATACGGCTAAATTTCAAT 1620
 Db 1608 gacctctcaatttctgacacttgataactattgctgctatacgcgttaataat 1667
 QY 1621 GGCTGTATGAGACAGGCTATTTCTTAACCGCTGACACCACTTGTAGCAATGGCTTACGGT 1680
 Db 1668 ggcctgtatgagacagctatcttaacgcttgcacacacacttgtagcaatggccttaacggt 1727
 QY 1681 CGCTGCCAAAGTTTACGCTCTATGATGACCTTCTGTGAAGAAACAATCCAGAGA 1740
 Db 1728 cgcgtccaaagtttaaccgctctatgatacctctgctgaaggaacatccagagga 1787
 QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGGGAATTTGTTAAATGGTGTCTCG 1800
 Db 1788 ttggaataatacagaggggctctacagaatgagaaatcgtatttaaaatggtgctcg 1847
 QY 1801 TTTCAGCTGGAACACTCTGCTCTACAAACACCCCATCAACTGAAGTTCAACCTCATC 1860
 Db 1848 ttctacgttgaagctcacactgctccacaacaccccccatcaactgaaagttcaagctc 1907
 QY 1861 ATCAGATAGTTCAACTTACAGTCTAGCTCAACCATTCACGAACAATAATATAGTACGAC 1920
 Db 1908 atcagatagttcaacttcaacagttctgaacacacccaagccaataatagtagac 1967
 QY 1921 TACCAATCCATACATTAATACGCAACAATCAAAATCAACCCCTGATACAAAATTCGAA 1980
 Db 1968 taccaatccataacataatgcaacaatacaaaacccctgatacaacaaatcagaa 2027
 QY 1981 TTCTCAACGACACAACA 1999
 Db 2028 tctcaacgacgacaaca 2046

RESULT 4
 AAH02055
 ID AAH02055 standard; DNA; 2160 BP.
 XX
 AC AAH02055;

XX 24-JUL-2001 (first entry)
 DT Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2048.
 XX
 DE
 XX Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitic;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN W0200123604-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-CA01150.
 XX
 PR 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 XX
 PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 XX
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX
 DR WPI; 2001-245006/25.
 XX
 PT Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 XX archaeal, bacterial, fungal and parasitic species in a test sample -
 XX
 PS Disclosure; Page 1463-1464; 1580pp; English.
 XX
 CC The present invention describes a method for generating a repertory of
 CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitic
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitic species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SO Sequence 2160 BP; 697 A; 522 C; 413 G; 528 T; 0 other;

Query Match 71.9%; Score 1438; DB 22; Length 2160;
 best Local Similarity 99.4%; Pred. No. 0;
 Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 1 TAAATCTACGACATTAATCAACTATGCTGAGCTTGGGTTCTGAGCGCCGTCGA 60
 Db 159 taaatctacgacataataatacaactatgctgacttgggttcgaaagcgcgca 218

QY	61	TGCCCAAGCTAAATGATATTTCCCAACATTTTGTTAAGGCAATCGTTTCTATCGAAGACCA	120
Db	219	tgcccaagctaaatgatatcccaacatttggttaaaggaaatcgcttcatcgaaagcca	278
QY	121	TGCGTTTGTGACACAGGGGGATGTATACCATTCGCTATCCGGGAGCTTTCCTCCGCA	180
Db	279	tgcgttcttcgacacaaggggagatgatatccacccgtaccgggagcttcttcgcaa	338
QY	181	TCGCAAAAGCAATTTCCCTCCAAAGTGGATCAACTCTCACCCCAACGTTGATTAACTTAC	240
Db	339	tcgcaagaagaattccctcccaagtgatcagctctcacccaacagttgattaagttgac	398
QY	241	TTACTTTTCACTTCGCATCTCCAGCACAAGCATTTTCTGTAAGGCTCAGAGACTGGTT	300
Db	399	ttacttttcaacttcgactctccgaccaagactattctcgtlaaggtccaggaagcttggtt	458
QY	301	AGGATTTAGTGTGAACAATAAAGCAACCAAGAGAATAATCTTGACTCTCTATATATAA	360
Db	459	aggatctcagttcgaacaaaaggaaccacaagaatacttcgacctctatataataa	518
QY	361	GGTCTACATGTCATATGGGAATGGAATGAGACAGCAGCTCAAACTACTATGTAA	420
Db	519	ggtctacatgtcctaalggaactatgaaatgagacagctcaaaactactatgtaa	578
QY	421	AGACCTCAATATTTTAAGTTTACTCTAGTTACCTTGCTGCTGGTAATGCTCAGACAC	480
Db	579	agacctcaaatlttaagtttactctcagttacgtcttgctgtaagtgcctcagcacc	638
QY	481	AAACCAATATGACCCCTATTCACATCACAAGACAGCCCAAGCCCGCAAACTTGCTTT	540
Db	639	aaaccaataatagccctattcacatccgaagcagcccaagacgcgcgaaaacttgctt	698
QY	541	ATCTGAATGAAAAATCAAGGCTACTCTCTGCTGAACAGTATGAGAAAGCAGTCAATAC	600
Db	699	atctgaatgaaaaatcaaggtctactctctctgtaacagtatggaagcagttcaatc	758
QY	601	ACCAATTACTGATGAGACTACAAAGTCTCAATTCAGCAAGTATTTACCTCGCTTACATGA	660
Db	759	accaattactgtagggcttaaaaagcttcaaatcagaagtaattacccttgcttaaatgta	818
QY	661	TAAATACCTCAAGGAAGTCTCATCAATCACTTGAGAAGAAAGGCTTAACCTTCATAC	720
Db	819	taattacctcaaggaagtcataccaagtctgaaagaagaacaaggtataaactctaac	878
QY	721	AACTGGAGTGGATGTCTACACAATGTAAACCAAGAGCTCAAAAACAATCTGTGGAGAT	780
Db	879	aactggagtgatgctctacccaatctgtagcaagaagtcataaactctgtagatc	938
QY	781	TTACATATACAGACGATATGCTTGCTTATCCAGACGATGATTTGCAAGTGCCTTACCAT	840
Db	939	ttacaatatcagaacgaatacgtctcctaccagaagatgtacagtcgcttctacatc	998
QY	841	TGTTGATGTTTCAACGGTAAAGTCAATTGCCCCAGGTAGAGAGCCCAATCGTCAAGTAA	900
Db	999	tgttgatgtttcaacggtaaaagtccttcgccaagcagaagagacgcacatcagtaagtaa	1058
QY	901	TGTTTCTCTTGGAATTTAACCAACACACTATGAACAAACCCGCACTGGGATCAACTATGA	960
Db	1059	tgtttctcttggaattaaaccaagcagtagaanaaaaacgcgcagctgggagtcatacatgaa	1118
QY	961	ACCGATCACAGACTATGCTCTCGCTTGGAAGTACGGTGTCTTAGATTCACTGCTACTAT	1020
Db	1119	accgatcacagactatgctctcgccttgtagtaagtgctctacaggtacaactgccaact	1178
QY	1021	CGTTTACAGTAGACCCCTATTAACACTACCCCTGGCAAAATCTCTGTTTAACTGGGATAG	1080
Db	1179	cgttcaagatgagccctataaactaacctcgggacaataacccttgcttataactggyatag	1238
QY	1081	GGGCTACTTTGGCAACATCACTTGGCAATACCCCTGCAACAAATGGCAAAAGTCCAGC	1140
Db	1239	gggctactttggcaacatacctccttgtaataagcctctgcaacaatcgcgaaaagctccagc	1298
QY	1141	CGTGAAGACTCTAAACAGGTGCGACTCAACCGCGCAAGACTTTCTAAATGCTATAGG	1200

Dd	1299	cg tgg a a a c t c t a a c a a g g t c t g a c t c a a c c g c g c c a a g a c t t c t c t a a t g t g t c t g g	1358
Oy	1201	a a t t c a a c t c c c c a a g a t a t t c a c t a c t c a a a n t g c c a a t t t t c a g t a a c a c a a c g a a t c a a	1260
Dd	1359	a a t g a c t a c c c c a a g t a t t c a c t a c t c a a a t g c a t t t c a a g t a a c a a c a a c c g a a t c a g a	1418
Oy	1261	c a a a a a a t a t g a g a c a a g t a g t g a a a a g a g g c t g c t g t a c g t t c c t t t g c a a a n t g	1320
Dd	1419	c a a a a a a t a t g a g a c a a g t a g t g a a a a g t g c t g c t g t t a c g t c t g c c t t g c a a t g g	1478
Oy	1321	t t g g a a c t t a c t a t t a a a c c a a t g t a t a t c a t a a g t c g t t t a g t a g a t g g g a g t a a a a	1380
Dd	1479	t g g a a c t a c t a t a a a c c a a t g t a t a t c c a t a a a g t c g t c t t a g t a g a t g g a t g a a a	1538
Oy	1381	a a g a g t t c t a a t g t c g a a c t c g t g c c a t g a a g a a g a a a g a c a c a g c c t a t a t g a t g a c c g a	1440
Dd	1539	a g a g t t c t a a t g t c g a a c c t c g t g c c a t g a a a g a a a g a a c a g c c t a t a t g a t g a c g a	1598
Oy	1441	c a t g a t a g a a a c a g t c t t t g a c t t a t g a a c t g a c g a a a t g c c t a t c t t t g c t g c t c c c	1500
Dd	1599	c a t g a t a g a a a c a g c t t t a g t a t g a a c t g a c g a a a t g c c a t c t t g c t a c t t g t g c t c c c	1658
Oy	1501	t c a g c c t g t a a a c a c g a a c c t a c a c t a c a c g a c g a a a a a t g a a a c a c a c a t c a a	1560
Dd	1659	t c a g g c t g t a a a a c a c g a a c c t c t a a c a t a c a g a c g a g a a a t g a a a c a a c a t c a a	1718
Oy	1561	g a c c t c t c a a t t t g t a g a c a c t g a t g a a c t a t t t t g c t g c t a t a c g t a a a t a t t a c a a t	1620
Dd	1719	g a c c t c t c a a t t t g t a g a c c c g t a g a a c a t a t t g t g t g t a t a c g c g t a a a t t c a a t	1778
Oy	1621	g g c t g t a t g a c a g c g t a t t t t a a c c g t c t g a c a c c a c t t t g a g s c a a t a g c c t t a c g g t	1680
Dd	1779	g g c t g t a t g a c a g c g t a t t c t a a c c g t c t g a c a c c a c t t g a g s c a a t a g c c t t a c g g t	1838
Oy	1681	c g c t c c c a a a g t t t a c c g t c t a t a t a t a t g a c a c t a c c t g t g t g a a g a a g a a t c c a a a a	1740
Dd	1839	c g c t c c a a a g t t a c c g t c t a t a t g a t g a c t a c t c t g t c t g a a g a a g a a t c c a a g a g a	1898
Oy	1741	t t g g a a t a t a c c a g a g g c c t c t a c a g a a t g a g a a t t c g t a t t t t a a a a t g t g t c t g c	1800
Dd	1899	t t g g a a t a t a c c a g a g g g c t c t a c a g a a a t g a g a a t t c g t a t t t a a a a t g g t g t c t g	1958
Oy	1801	t t t c a t a g t g g a a c t a c t g t c t c a c a c a c a c c c c a t c a a c t a g a a a g t t c a a g c t a t c	1860
Dd	1959	t t c t a c t g t g a g c t c a c c t g t c c a a c a a c c c c a t c a a c t g a a a g t t c a a g t c a a g t c a t c	2018
Oy	1861	a t t c a a t a g t t a a c t t c a c a c t t r a g c t t c a a c c a c t c c a a c a c a a t a t a t a t a t a g t a c g a c	1920
Dd	2019	a t c a a g a t a g t a a c t c a a c t c a g t c t a g c t c a a c c a c t c c a a g c a c a a t a a t a g t a g c a c	2078
Oy	1921	t t a c c a t c c t a c a a t a t a t a t a c c a a c a a t a t a c a a t a c c c t g a t c a a c a a a t c a g a a	1980
Dd	2079	t a c c a a t c c t a a c a a t a a t a g c a a c a a t c a a a t a c a a c c c t g t a c a a c a a a t c a g a a	2138
Oy	1981	t c c t c a a c c a g c a c a a c c a 1999	
Dd	2139	t c c t c a a c c a g c a c a a c c a 2157	
RESULT	5		
ID	AAS55688	standard; DNA; 2160 BP.	
XX	AAS55688;		
AC			
XX			
DT	13-FEB-2002	(first entry)	
XX			
DE	Streptococcus pneumoniae DNA for cellular proliferation protein #259.		
KM	Antisense; ds; prokaryotic cellular proliferation gene;		
KM	antibiotic; antibacterial; drug design.		
XX			

OS Streptococcus pneumoniae.
XX
PN W0200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI: 2001-611495/70.
DR P-PDB: AAU37829.
XX
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
PS Claim 27; Seq ID No 9325; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence encodes an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 2160 BP; 697 A; 522 C; 413 G; 528 T; 0 other;

Query Match 71.9%; Score 1438; DB 23; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATATAATCACTATCTGCTGAGCTTGGCTTCTGAGCCGCTCAA 60
Db 159 taaatctagacaataataataatcaactatctgacttggtctgaagccgcgctcaa 218
QY 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGCAATCGTTTCTATGAAAGCA 120
Db 219 tgcaccaagctaataatattcccaagatttggttaagcaatcgtttctatcgaaagca 278
QY 121 TCGCTTCTTGACACACAGGGGGATTTGATACATCCGATCTCTGGAGCTTTCTTGCCAA 180
Db 279 tcgcttcttgaccacacaggggattgataccatccglatcctggagcttcttgcgaa 338
QY 181 TCGCAAGCAATTCCTCCCAAGGTGATCAACTCCACCAACAGTTGATTAAGTTGAC 240
Db 339 tcgcaagaacatctccctcaaggtgataagctctcaactcaaaagttgataagttgac 398
QY 241 TTACTTTCACTTGCAGCTTCGACGACGACTATTTCTGTAAGGCTCAGGAAGCTTGTT 300

Db 399 ttactttcaacttcgacttcgcagcagactatttcctgtaaggctcaggaagcttgctc 458
QY 301 AGCATTCAGTTAGACAAAAAGCAACCAAGAAATCTTGACCTACTATATAATTA 360
Db 459 agcatctcagttagaacaaaaaagaacaaagcaagaatcttgactactataataaa 518
QY 361 GGCTACATGCTAATGGAACCTATGGAATGCACAGCAGCTCAAAATCACTATGTGTA 420
Db 519 ggtctacatgtctaatggaacctatggaatgagcagcgcgtctaaactactctgttaa 578
QY 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGCTGGTGAATGCCCTGAGCAC 480
Db 579 agacctcaataatttaagtttacctcagttagccttgctgctgtaagtgcctcagcacc 638
QY 481 AAACCAATGAGCCCTATTCACATCCAGAAAGCCCAAGACCGCGGAACCTGGCTT 540
Db 639 aaaccaataatgacccctattcaacatccagaagcagcccaagaccgaaacttgctc 698
QY 541 ATCTGAATGAAAAAATCAAGGCTACATCTGCTGTAACAGATATGAAAGCAGTCAATAC 600
Db 699 atctgaaatgaaaaaatacaaggctacatctctgcctgaaacagtatagaagcaatcac 758
QY 601 ACCAATTAATGAGTACACAAAGCTCAAAATCAGCAAGTAAATTAACCTGCTTAATGA 660
Db 759 accaattactgattggtctacaagaagctcacaatcagcaagtaattaccctgcttaca 818
QY 661 TAAATTAACCAAGGAGGATCAATCAAGTTGAAGAAAGAAACAGCTATTAACCTACAC 720
Db 819 taattaccatgaagaagatcaacatcaagttgagaagaacaagcattacaactaccac 878
QY 721 AACTGGATGATGATCTTCACAAATGTAGACCAAGAAAGCTCAAAACATCTGTGGATAT 780
Db 879 aactggatgattgattctacaacaaatgtagccaagaagctcacaacatctgtggat 938
QY 781 TTACAAATACAGACAAATACGTTGCTTATCCAGACGATGAATTTGCAATGCTTACCAT 840
Db 939 ttaacaatacagacgaatcagcttgctctaccagatgaattcgaagctccttaccat 998
QY 841 TGTGATGTTTCTTAACGCTTAAGTCATTTGCCACCTTGAGAGACAGCATCAAGTAAT 900
Db 999 tgtgattgttcttaagtgtaagatcattgctccagctgagagcagccatcaagttaa 1058
QY 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAACAACCCGCACTGGGATCAACTATGAA 960
Db 1059 tgttccttggaaatlaaccaagcagtagaacaacccgcagctgggataactatgaa 1118
QY 961 ACCGATACAGACTATGCTGCTGCTTGAGATACGATGCTACAGATTCACATGCTACTAT 1020
Db 1119 accgatcacagactatgctcctgcttgtagttagcgttcaagtaacttcgactat 1178
QY 1021 CGTTACAGATGAGCCCTATATACTACCTCGGACAAATATCTTGTATTAAGTGGATAG 1080
Db 1179 cgttacaagatgagccctataactacccctgtagaacaataaccctgttataacttgg 1238
QY 1081 GGGCTACTTTGGCAACATCACCTTGCAATACGCCCTGCAACATCGGAAACGTCCAGC 1140
Db 1239 gggctacttggcaacaatacacttgcaatagccctgcaacaatcggaaacgctccagc 1298
QY 1141 CGTGAATCTTAACAGAGTCCGAGCTCAACCGCGCAAGACTTTCTTAATGTGTAGG 1200
Db 1299 cgtgaaactctaaacaagtcggaacccaacgscgcaagaacttctctaaatgtctcg 1358
QY 1201 AATGACTATCCCAAGTATTCATCTCAATTCGCAATTCGCAATTCGCAATTCGCAAT 1260
Db 1359 aatgactatcccaagatltcaactcaactcaaatgcaattcaagtaacaacaacgcaatcaga 1418
QY 1261 CAAATAATATGAGCAGTAGTGAAGATGGCTGCTTACGCTCTCTTGAAGATGG 1320
Db 1419 caaataataatgagcagtagtgaagaatgagctgcttctttagcgtcctttagaaatg 1478
QY 1321 TGAAGTTACTATTAACCAATGATATATCAATAAAGTGTCTTTAGATGGAGATGAAA 1380
Db 1479 tgaagttactatataaccaaataatgatatcaataaagtgctcttttagatggagatgaaa 1538

Db 1479 tggaaactactataaacaatgtatataccataaagtcgctttagtbgatggagtgaa 1538
QY 1381 AAGATTCTATATGCGAAGCTCGTCCATGAGGAAGAGAGCTATATATACCGA 1440
Db 1539 aagatctctcaatgcggaacccgctgcaatgaaagaagaacgactatagatgacga 1598
QY 1441 CATGATGAAACAGTCTGACTTATGAACTGAGCAAAATGCTTCTGTTGCTCC 1500
Db 1599 catgtatgaaaaacgctctgagatctgaaactgagcaaaatgcctatctgtctcc 1658
QY 1501 TCAGGCTGTAAACAGAGAACTTACTATACAGACGAGAAATGAAACACATCA 1560
Db 1659 taagctggtaaaacaggaaccttaactacagacgaggaatgtgaaacacataca 1718
QY 1561 GACCTCTCAATTGTAGACCCGATGAACTATTTGCTGCTTACGCCGTAATTCAT 1620
Db 1719 gacccctcaattctgagcccgagaaactctgtgctatagcgaataatcaat 1778
QY 1621 GCGTGTATGACAGGCTATTTTACCGTCTGACACCACTTGTAGGCAATGCGTTACGGT 1680
Db 1779 ggcgtatgagacaggtatcttaaccgctctgacacactgtgagcaatgagccttaag 1838
QY 1681 CGCTGCAAAAGTTTACCGCTCTATGATGACCTACTGTCTGAGAGCAATCCAGAGA 1740
Db 1839 cgcgtccaaagttacacgctctatgataagcctactgcttgaggaaggaatccagaga 1898
QY 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTAATAAATGGTCTCG 1800
Db 1899 ttggaatatacagaggggctctacagaaatgagaaatcgatcttaaaatggtctcg 1958
QY 1801 TTCTAGTGAAGTCACTGCTGCTCCACAAACCCCATCACTGAAAGTTCAAGTCAATC 1860
Db 1959 ttctacgtaggagctactcgtcccaacaacccccatcaactgtgaagtcaagctcacc 2018
QY 1861 ATCAGATAGTTTCACTTCACTGACTGACTCAACCACTCCAGACAAATATATGACGAC 1920
Db 2019 atcagatagttcaactcaactcagctagctcaaccactccaagcacaataatagtagc 2078
QY 1921 TACCAATCTCAATATATAGCAACATCAATCAACCCCTGATCAACAATAATCAGAA 1980
Db 2079 taaccaatccatacaataacgcaacaatacaatacaaccctgatacaacaataatcaga 2138
QY 1981 TCCCTACCCAGCAACCA 1999
Db 2139 tccctaacacagcaacaacca 2157

RESULT 6
AAH02067
ID AAH02067 standard; DNA; 1249 BP.
XX
AC AAH02067;
XX
DT 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SPO ID NO:2060.
XX
KW Species specific; genus specific; family specific; probe; detection;
KW Identification; algal; archaeal; bacterial; fungal; parasitica;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN MO200123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000MO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.

PR 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
DR WPI, 2001-245006/25.
XX
PT Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitica species in a test sample -
PS Claim 27; Page 1473; 1580pp; English.
XX
CC The present invention describes a method for generating a repertory of
CC nucleic acids of tut, tus, atpd and/or reca genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitica
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitica species, genus, family and group. A nucleic acid (1)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SO Sequence 1249 BP; 402 A; 302 C; 246 G; 299 T; 0 other;

Query Match 59.9%; Score 1198; DB 22; Length 1249;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 720 CAACGTGGATGATGCTTACACAAATGTAGACCAAGCTCAAAACATCTGTGGATA 779
Db 1 caactggagatgagtctcacacaatgttagaccaagaagctcaaaaactcgttggata 60
QY 780 TTTACAATACAGAGCAATAGTGTGCTTATCAGACGATGAATGTGCAAGTGGTTCTACCA 839
Db 61 tttaacaatacagaagatacgttgcctatccagacagatgaattgcaagtcgcttcaacca 120
QY 840 TTGTTGATGTTTCTAACGCTTAAGTCATTTGCCCACTTAGACACAGCCCTACGTCAAGTA 899
Db 121 ttgttgatgttcttaacggtlaaagtcattgcccagctgagagcagccatcagtaagta 180
QY 900 ATGTTTCCCTTGGCAATTAACCAAGCAGTAGAACAACCCGCACTGGGAGATCAACATGA 959
Db 181 atgttctcttggaaatlaaccagaagtagaacaacaacgcgacttgggagataactatga 240
QY 960 AACGATACAGACATATGCTCTGCTTGGCTTGGATAGCGTGTCTACGATTCACCTGTACTTA 1019
Db 241 aacgatacagaactatgctcctgcttggagtagtgcgttcaagatcaactgataccta 300
QY 1020 TCGTTCACGATAGCCCTATATACCTCGGACAAATATCTCTTTATATCTGGGATA 1079
Db 301 tcgttacagatgagccctataactaccctgggagacaatactctcgttataacttgggata 360
QY 1080 GGGGCTACTTTGGCAACATCACCTTGCAATATAGCGCCCTGCACAAATCGGGAACGTCCAG 1139

D	b	361	ggggctacttttggaacatacacttgcgaatcagcccttgaaacaatcggaaacgtccag	420			
Q	y	1140	CCGTGGAAACTCTTAACNAGGTGGACTACCGCGGCCCAAGCTTTCTTAATGGTCTAG	11399			
D	b	421	ccgtggaactctaaacaagaagtcgagctcaacccgcgaagactctcctaattgctctag	480			
Q	y	1200	GAATCGACTACCCNAGTATTCACCTACTCTCAATCCATTTTCAGTAAACACACCGAATCAG	1259			
D	b	481	gaatcgactaaccgaagatctcactactcaaatgcgaatcttcaagtacaacacagcatag	540			
Q	y	1260	ACAAAAAATATGAGCAGTAGTGTAAAAAGATGGCTGCTCTTAGCGCTGGCTTTGGCAATG	1319			
D	b	541	acaaaaaatcttgagcaagtagtgaaaagaatgctgccttgctctcagctgccttgcaatg	600			
Q	y	1320	GTGAACTTCTACTTAAACCAATGATATCCATTAATGGCTCTTTGTGTATGGAGAGTAA	1379			
D	b	601	gtggaactctactataaaccaatgataccataaagtgcctcttgatgaggaatcgaa	660			
Q	y	1380	AAGAGTTCTCTAATGTGTGGAACCTCGTCCATGGAAGAAAGCAGACCTATATGATGACCG	1439			
D	b	661	aagagctctctaatgctcgaagctcgcgcgaatgaagaaagcagcctatattgatgcag	720			
Q	y	1440	ACATGATGAAAAACAGCTTGTACTTATGGAACAGGAGAAATCCATATCTGTTGGCTCC	1499			
D	b	721	acatgatagaacaacgctcttgaactctatggaactcgtgaagaaatgcctatcttgctgctc	780			
Q	y	1500	CTCAGCGTGTGTAAACAGAAACCTCTTACTATACAGACGAGAAATTTGAAAAACCATCA	1559			
D	b	781	ctcagctctgttlaaaaaacgagacctcttaactatacagaagaggaatctgaaaccacatca	840			
Q	y	1560	AGACCTCTCAATTTGTATGACACCTGATGAACTATTGCTGGCTATACGGGTAAATTTTAA	1619			
D	b	841	agacctctcaatttgttagaccctgtaagaaactattgtgcgcatacgcgtlaaataatcaa	900			
Q	y	1620	TGGCTGTATGTAGACAGCGTATTTCAACCGTGTGACACCACTTTGAGCAATGGCTTACGG	1679			
D	b	901	tggctgtatgtagcaagcgctatcttaaccgcgtcgaacacactctgttagcaatgacctag	960			
Q	y	1680	TCGCTGCCAAGTTTACCGCTCTATGATGACCTACTGTCTGAGAGGAACAAATCCAGAG	1739			
D	b	961	tcgctgcgcaaaagtttacccgtctatgatbaacctactcgtctctaaagaaatccaaga	1020			
Q	y	1740	ATGTGAATATATACGAGGGGCTCTACAGAAATGGAATTCGTATTAAAAATGGTGTCTC	1799			
D	b	1021	atgtgaataataccgaagggcctcaagaagaatggaatcttgatcttaaaaatgctgtctc	1080			
Q	y	1800	GTTCTACGTGAACCTCACCTGCTCCACAAACACCCCATCAACTGAAAGTTCCAACTCAT	1859			
D	b	1081	gttctacgttagagctcacctgcctccacaacaaccccccatcaactgaagtccaagctcat	1140			
Q	y	1860	CATCAGATATTTCAACTTCACACTCTAGACTCAACCACTCCAGCAACAAATATATGATAGA	1919			
D	b	1141	catacagtagtcaactctcaactcagctcagctctcaaccctcaagcaacaataatagtacga	1200			
Q	y	1920	CTACCAATCTTAACAATATATACCAACATTCAAATTCACACCCCTGATGA	1968			
D	b	1201	ctaccaatccttaacaataataatagcacaacatcaatacaacccttgaatca	1249			
RESULT 7							
AAZ35952 standard; DNA; 1260 BP.							
AC	AAZ35952:						
XX	07-FEB-2000	(first entry)					
DE	Streptococcus pneumoniae pbp1A TER isolate n) nucleotide sequence.						
KW	Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;						
KW	transpeptidase encoding region; TER; antibiotic resistance; diagnosis;						
KW	transcription; identification; pneumococcal meningitis; ss.						

```

Db 403 gtccctgacctgagtagtggtctacgaltcaactgctactatcglttcagatgagccc 462
Oy 1037 TATTAATACCGGGGCAAAATCTCTGTTTAACTGGGATAGGGGCTACTTTGGCAAC 1096
Db 463 tataatacccccggagcaaatactccgtgttataacttggaatagggcactcttggaac 522
Oy 1097 ATACCTTGCAATAGAGCCCTGCAACATCGCAAAAGTCCAGCCGTGGAACTGTAAAC 1156
Db 523 atcactctgcaatacgcctctgcaacaatcgcaaacgtcccgcttggaactcttaac 582
Oy 1157 AAGGTGGAGCTAACCCGGCCAAAGACTTTCTTAATGCTTGAAGATGACTACCCAACT 1216
Db 583 aagtgtagcctaacgcgcgaagcttccctaaatggtctaggaatcgactaccaagt 642
Oy 1217 ATTCACTACTAAATGCCCTTTCAAGTATACACACGAAATCAGACAAAATATGAGACA 1276
Db 643 attcaactactaaatgaccttcaagtaaacacaccgaatcgacacaaaatattgagca 702
Oy 1277 ACTAGTGAAGATGGCTGCTGCTTACGCTTGCAATGCGTGAAGTGAAGTACTATATA 1336
Db 703 agtagtgaaaagatggtcgtctctacgtctgaccttgcaaatggtggaactactataaa 762
Oy 1337 CCAATGTATATCCATAAAGTCTTAACTGATGAGAGTGAAGAAAGTCTCTAATGTC 1396
Db 763 ccaatgtatataccataaagctgctttagtgatgagtgagaaagagcttctaattgctc 822
Oy 1397 GGAACCTCGCCCATAGAGAAAGACAGACGCTTATGATATACCGACATGATGAAAACAGTC 1456
Db 823 ggaacctcgccatgaaagaaagacagcctatatgataccgacatgattgaaacagtc 882
Oy 1457 TTGACTTATGAACTGAGAAATGCTATCTTGGCTTGCTCCCTCAGAGCTGTAAACA 1516
Db 883 ttgagttatggaactcgagcaaatgctctatctgtcttgctccctcaggtctgtaaaaa 942
Oy 1517 GGAACCTCTAATATATACAGACAGAAATGTAACACACATGAAAGCTCTCAATTTGTA 1576
Db 943 ggaacctcaactatacagacgaggaatgaaacacacatcaagaacctctcaattgta 1002
Oy 1577 GCACGTGATGAACTATTTGCTGCTATACGCGTAAATATTCATGAGGTATGACGAGCG 1636
Db 1003 gcaactgaagactattgctggtcctatacagcgtaaatatcaatgctgtaaggcaagc 1062
Oy 1637 TATTTAAACGCTGTGACACACACTTGTAGCAATGGCTTACGGCTGCTGCCAAAGTTTAC 1696
Db 1063 tatttaacocgtctgacacaccttgtagcaatggtccttaagctgctgcaagaattac 1122
Oy 1697 CGCTCTATGATACCTACTGCTGTGAAGAGCAATCCAGAAATTTGGATATACCAAG 1756
Db 1123 cgctctatgatacctactctgtaaggaagcaatccagagagattggaatataccagag 1182
Oy 1757 GGGCTCTACAGAAATGAGAAATTCGATTTAAATGGTGGTTCGTTACGCGAAGTCA 1816
Db 1183 gggctctacagaatggaagaaattcgatttaaaaatggtccgttctacgtcgagactca 1242
Oy 1817 CCTGCTCCACAACAACCC 1834
Db 1243 cctgctccaacaacacc 1260

```

```

RESULT 8
AAZ35953
ID AAZ35953 standard; DNA; 1260 BP.
XX
AC AAZ35953;
XX
DI 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1A TER isolate o) nucleotide sequence.
XX
KW Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
KM transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KM detection; identification; pneumococcal meningitis; ss.

```

```

XX
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
XX
PA (UWI-) UNIV WITWATERSRAND.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX
DR WPI; 1999-601770/51.
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
XX
PS Claim 11: Fig 4; 63pp; English.
XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 405 A; 293 C; 260 G; 302 T; 0 other:

```

```

Query Match 47.0%; Score 940; DB 20; Length 1260;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1190; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Oy 617 CTCAAAAGTCTCAATCAAGCAAGTATTAACCTGTTACATGATTAATTACCTCAAGGAA 676
Db 43 ctcaaaagtcctcaaatcagcaagtaattacccctgttacaatgataattactcaagaa 102
Oy 677 GTTATCAATCAAGTTGAAACAGAAACAGGCTTAACTTACTACAACTGGATGATGTC 736
Db 103 gttatcaatcaagttgaaacagaaacaggtctaaacctactcaacaactggaatgtaatg 162
Oy 737 TACACAAATGTAGACCAAGCAAGCTCAAAAACATCTGTGGGATATTATCAATACAGCA 796
Db 163 tacacaatgttagacccaagaagatcaaaaaacatctgtyggatatttacaatacagaagaa 222
Oy 797 TACGTTGCCCTATCAGACGATGAATTCGAAGTGGCTTACCATGTTGATGTTTCTAAC 856
Db 223 tacgttgccctatccagagatgaattgcaagtcgcttaccatgttgatgtttctaac 282
Oy 857 GGTAAAGTCATTGGCCAGCTAGAGACAGGCCATGAGTCAAGTAATGTTTCCCTTGGAAAT 916
Db 283 ggtaaagtcattggccagctagagacagccatcagtaagaatggttccctcggaatt 342
Oy 917 AACCAAGAGTAGAACAACCGCAGCTGGGATGATCAATGAAACCGTACAGACTAT 976
Db 343 aaccaagcagtagaacaacacgcgacgtgggagtaactatgaacccgatacagaactat 402
Oy 977 GCTCTGCTTGGCTGGAATGAGTGTCTACGATTCACCTGACTATATGCTTACAGATGAGCC 1036

```

Db 403 gctccgctctggaagtcggtctacgaltcaactcgtctactcgtctcaagtagagcc 462
QY 1037 TATAACACCTGGAGCAATACCTGCTTTATTAACGTGGATAGGGCTCTTTGGGCAAC 1096
Db 463 tataactacccttggagaataatccccgtttataactcgtgataagggtactcttggcaac 522
QY 1097 ATCACTTGAATACGCCCTGCAACAATGCGGAACGTCGCCAGCCGTGGAACCTTAAC 1156
Db 523 atcaccttgaataacgcccctgcacaacatcgcaaacgttcccaagcctgggaaccttaac 582
QY 1157 AAGTGGAGTCAACCGCCGCAAGACTTCTTAATGCTTAGAATACGCTACCCAGT 1216
Db 583 aagtcgactcaacccgcgaagacttccctaagtcgtaacgtaacactaccag 642
QY 1217 ATTCACTACTCAATGCGCATTTCAAGTACACAACCGAATCAGACAAATAATAGAGCA 1276
Db 643 attcacactcaaatgcatcattcaagtaacacaacacgaatcagcaaaaataatggagca 702
QY 1277 AGTAGTAAAAAGATGCTGCTGCTTACGCTGCTTTGCAANTGCTGGAACCTTACTATAA 1336
Db 703 agtagtgaagaagatgctgctgcttactacgctccttgcaaatggtggaaactactataa 762
QY 1337 CCATGATATTCATTAAGTCCTTTAGTATGGAGTGAAGTGAAGAGTCTCTAATGTC 1396
Db 763 ccaatgtatatacctaagaagtcgtctttagtgatggagtgaaaaagatctctaatgctc 822
QY 1397 GGAATCTGTCATGAAGAAAGACAGACCTATATGATGACCGACATGATGAAGAAACAGTC 1456
Db 823 ggaactctgcatgaagaagaagacagcctatatatgtagccgacatgatagaaacagtc 882
QY 1457 TTGACTTATGAACTGAGCAAAATGCTTATCTGCTTGGCTCCCTCAGGCTGTAATAACA 1516
Db 883 ttgacttatgaaactgagcaaaatgcttacttctgcttgcctcagctcgtgtaaaaaa 942
QY 1517 GGAACCTCTACATATAGACGAGGAATGAAATGAAACCAATCAAGACTCTCAATTTGTA 1576
Db 943 ggaacctcaactatacagaaggaagaatgaaataccacatcaagaccttcaattgta 1002
QY 1577 GCACCTGATGAACTATTTGCTGCTATACCGCTTAATATTTCAATGCTGATGAGACAGC 1636
Db 1003 gcaacctgaatgaactatctgtgctatacgcgtataatattcaatgctgtatgagacgyc 1062
QY 1637 TATTTTAACTGCTGACACCACTTTGATGAGCAATGCGCTTACGGCTGCCAAAGTTTAC 1696
Db 1063 tatctcaacgctcgcacacactttagagcaatgagccttaagtcgctgcacaaagttac 1122
QY 1697 CGCTCTATGATGACCTACCTGCTGGAAGGAAGCAATCCAAAGATTGGAATATACAGAG 1756
Db 1123 cgcctctatgatgacctcctgctcgaaggaagcaatccagaatgtgaaatataccagag 1182
QY 1757 GGGCTCTACGAATGAGAAATCGTATTTAAATAGTGCCTGCTTAGCTGGA 1811
Db 1183 gggctctacagaatgagaaatcgtaattcaataaagtgctcgtcttcaagctgga 1237

RESULT 9
AAZ35950
ID AAZ35950 standard; DNA: 1260 BP.
XX
AC AAZ35950;
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1a TER isolate 1) nucleotide sequence.
XX
KW Streptococcus pneumoniae; penicillin binding protein: pbp2b; pbp1a;
KM transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
XX detection; identification; pneumococcal meningitis; ss.
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.

XX 28-APR-1999.
PD 05-AUG-1998; 982A-0007024.
XX
PF 01-AUG-1997; 972A-0006886.
XX
PR (SAME-) SOUTH AFRICAN INST MEDICAL RES.
XX (UYWI-) UNIV WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX WPI: 1999-601770/51.
DR
XX Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX
XX Claim 11; Fig 4; 63pp; English.
PS
XX A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2b (pbp2b) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC identified for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1a transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 407 A; 293 C; 258 G; 302 T; 0 other;

Query Match 45.2%; Score 903; DB 20; Length 1260;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 575 GAACAGTATGAGAAAGCAGTCAATACCAATTAATGATGAGTCAATCAATCA 634
Db 1 gaacagtatgagaaagcagtcacatcaacatctactgtagcacaagaatctcaatca 60
QY 635 GCAAGTATTAACCTGCTTACATGATGATTAATTAATTAATTAATTAATTAATTA 694
Db 61 gcaagtaattacacctgcttcaatgataatcaatcaagaagatcaatcaatcaagttgaa 120
QY 695 GAAGAAACAGGCTTATACCTACTACACACTGGATGATGATGATGATGATGATGATG 754
Db 121 aaagaaacaggtctataaccgctcaacacggtgagatgtctacacaaatgtgagacaa 180
QY 755 GAAGCTCAAAAACATCTGCTGATATTTATCAATACAGCAAGATACGTTGCTATCCAGAC 814
Db 181 gaagctcaaaaacatctgctgagatattcaatacaagaagaatcgttgcataccagac 240
QY 815 GATGAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 874
Db 241 gatgaattgaagtcgcttcaatcattgttgcattgttcttaagcgttaagtcattgcccag 300
QY 875 CTAGAGACAGCCATCATGTAATGTTTCTTCCGAAATTAACAGCAAGTGAAGAAACA 934
Db 301 ctgagagacagccatcagtcagtaagtaagttcttccctcggaatlaaccaagaagtagaaca 360
QY 935 AACCGGACATGGGAGATCAATGAAACGATCAACACATGATGCTGCTGCTGGAATAC 994
Db 361 aaccggacatgggagatcaatgaaacgatcaacacatgattgctcctcgttggagtagc 420
QY 995 GGTGCTACGATTAACCTGCTACTATGCTTACAGATGAGCCCTATAACTACCTCGGACA 1054

```

Db 421 ggtgtctacgattcaactcgtactatcgttcaagatgagccctatactacccctggaca 480
QY 1055 AATACTCCTGTTATTAATGAGGATAGGGCTACTTTGGCAACATCACTTGCANATCGCC 1114
Db 481 aataccctcgttatatacagcggatagggcctacttgcatacaataccttgcatacagcc 540
QY 1115 CTGCAACATCGGAAAGCTCCAGCGGTGGAACCTTAACAGGTGCGACTCAACCGC 1174
Db 541 ctgcaacaacgcgaaacgtcccaagcgtggaacctaacaagctcgagcaacgcgc 600
QY 1175 GCCAAGACTTTCTTAATGATGAGTACTAGCACTACCAAGTATTCATCAATAGCC 1234
Db 601 gccaaagacttctctaataigtctcggaatcgactaccgaagtatcactacccaatgccc 660
QY 1235 ATTTCAAGTAAACACCAATCAACAAATAATGAGCAAGTATGTAAGTAAAGATGCT 1294
Db 661 attcaagtaacacacacgaatcagacaaaataatgagcaagtagtgaagaatgagct 720
QY 1295 GGTGCTTACGCTGCTTTCATATGATGAGTACTTACTATTAACCAATGATATCCATAA 1354
Db 721 gctgtctaacgctgcttgcgaatggtgaaactactataaaccacgaatcattccataaa 780
QY 1355 GTGCTCTTATGATGAGGAGTGAAGAAAGATTCTTAATGTCGGAATCGTGCATGAAG 1414
Db 781 gtcgtctttagtgatgagggatgaaaagagctctcaatgctcggaactcgctcgaatgag 840
QY 1415 GAAAGACAGCCTTATGATGATGACCGCATGATGAAAGAGTCTTGACTTATGCACTGGA 1474
Db 841 gaaacacacgcctataatgatacgcgacatgataaagacgtcttgagtagtaactgga 900
QY 1475 CCAAAATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1534
Db 901 cgaatgctcatcttgcgtgctgctcctcagctggtgtaaaacggaaacctatacaca 960
QY 1535 GACGAGGAATGAAACACCATCAAGACCTCTCAATTTGTAGCACCTGATGAATATT 1594
Db 961 gacgaggaatgaaaacacacatcaagacctctcaattgttagcacctgacgaactatt 1020
QY 1595 GGTGCTATAGCGCTAAATATTCATGCTGATGAGACGCTATTTCTTAACCTGAGACA 1654
Db 1021 gctgtgtctaaacgcgtataaattcaatgctgatacgcagagctatctaccgctcgaca 1080
QY 1655 CCACCTGTATGAGCAATGCGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1714
Db 1081 ccactttagtgcaatgagccttactgctgctgctgcaaaagttaacgcctctatgtagctac 1140
QY 1715 CTGCTGAAAGAGCAATCCAGATGTTGGAATATPACAGAGGGGCTCTACAGAAATGGA 1774
Db 1141 ctgtctgaaggaagcaatccagagatgtaataatacagaggggctctacagaaatgga 1200
QY 1775 GAATTCGTATTTAAATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1834
Db 1201 gaatcgtatatttaaaaatggtgctgctctactgtaactcactgctcacaacaaccc 1260

```

RESULT 10
AA235951
ID AA235951 standard: DNA: 1260 BP.

AC AA235951;
DT 07-FEB-2000 (first entry)
XX Streptococcus pneumoniae pbp1A TER isolate m) nucleotide sequence.
DE Streptococcus pneumoniae: penicillin binding protein: pbp2B; pbp1A;
KW transpeptidase encoding region; TER: antibiotic resistance; diagnosis;
KW detection; identification; pneumococcal meningitis; ss.
XX Streptococcus pneumoniae.
OS
XX
PN ZA9807024-A.

```

XX 28-APR-1999.
PD 05-AUG-1998; 982A-0007024.
PF 01-AUG-1997; 97ZA-0006886.
PR (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UTWI-) UNIV WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
PI Klugman KP, Smith AM, Du Plessis M;
DR WPI: 1999-601770/51.
XX Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
PS Claim 11; Fig 4; 63pp; English.
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC S. pneumoniae can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
SQ Sequence 1260 BP: 407 A; 293 C; 258 G; 302 T; 0 other:

```

Query Match 45.2%; Score 903; DB 20; Length 1260;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 575 GAACGATATGAGAAAGCACTCAATACCAATTAATGATGAGTACTCAAGTCTCAATCA 634
Db 1 gacagatagagaagaacagatcaatacacaatactgagactacaaagctccaatca 60
QY 635 GCAAGTATTTACCTGCTTACATGATGATATTAATTAATTAATTAATTAATTAATTAAT 694
Db 61 gcaagtaattaccctcgtctactatgataattacccaagaagatcacaatcaagttgaa 120
QY 695 GAAGAAACAGGCTATTAACCTACTCAACACTGGAGATGATGCTTACCAAAATGTAGCCAA 754
Db 121 aaagaacacgctataaactgctcaaacctggatgaltgctacacaagtgaagccaa 180
QY 755 GAAGCTCAAAAACATCTGTGGATATTTTACATATACAGAGCAATAGCTTCCATCCAGAC 814
Db 181 gaagctcaaaaacatctgtggatatttacaatacagagaatagctgtccatccaagac 240
QY 815 GATGAATTCAGATGCTGCTTACCAATTTGATGTTTCTTAACGGTAAATCAATTTGCCAG 874
Db 241 gatgaattgcaagtgcctctacccattgttgtatgtttcctcaagygtaaatgcatcgcag 300
QY 875 CTAGAGACAGCCCATCAGTCAAGTAAATGTTTCTCGAATTAACCAACAGCAGTAGAACA 934
Db 301 cttagagacagccatcagcaagtaagtgttccttcogagataacacaagcagtagaaca 360
QY 935 AACCCGATGTTGGGATCACTATGAACCGATACAGATATGATGCTTCCCTTGGATGAC 994
Db 361 aaccgcgactgggatacactatgaaacgcgatcaagatacgtctcgtcgttagtacc 420
QY 995 GGTGCTACGATTCAGTCTACTATCGTTACGATGAGCCCTATTAATCACTGAGACA 1054

```



```

Db      421  ggtgtctacgattcaactcgtctacatcgtctcagcatgagccctaaactcctggagaca 480
OY      1055  AATATCTCTGTTTATATACCTGGATGGGGCTACTTTGGCAACATCACCTTGCAATACGCC 1114
Db      481  aatactcctgtctataactcgtggaagaggtcactcttgcaacacacacctcttgcaatagcc 540
OY      1115  CTGCACAAATCGGAAACGTCGCCAGCCGTGGAACCTTAAACAGATCGGACTCAACGCC 1174
Db      541  ctgcacaacatcggaacacgtccacgctggaactcttaacaaggtccggaactcaacgcg 600
OY      1175  GCCAAGACTTTCTCTAAATGCTCTAGAGATGACTACCCAGATTTACTACTCAATATGCC 1234
Db      601  gccaaagacttccctaaatggtctcggaatcgactacccaagttactcactcctcaaatgccc 660
OY      1235  ATTTTCAGTAAACACCAACGGAATCAGACAAAATATGAGCAAGTACTGTAAGAAATGGCT 1294
Db      661  atttcaagtaacacacacggaatcagacaaaaataatagaaagtagtgaagaatgagct 720
OY      1295  GCTGCTTACGCTGCTTTGGAAATGCTGGAACCTTACTATTAACCAATGTATATCCATAAA 1354
Db      721  gctgcttaacgctgctcttgtaaaatggtggaacttactataacccaatgtatattcctataa 780
OY      1355  GTGCTCTTTTGTGATGGAGTGAAGAAAGAGTTCTTAAATGTCGGAATCTGTCATGAAAG 1414
Db      781  gtctgtctttagttagtgagagtgaaagagttcttctaattgctggaactcgtgcatgag 840
OY      1415  GAAACGACAGCTTATATGATGAGACCGACATGATGTAAGAAACAGCTTGACTTATGGAACG 1474
Db      841  gaaacacacgctataatgatacgacacatgataagaaacagcttgagttatggaactcgga 900
OY      1475  CGAAATGCTATCTTGTGCTTGGCTCCCTCAGGCTGTGTAACAGGAACTTACTATATACA 1534
Db      901  cgaatgctctatctgttgcttgctccctcaggtcgtgtaaaacaggaactcttaactataca 960
OY      1535  GACGAGGAATTTGAAACCCACATCAACACCTCTCAATTTGTAGACCTGATGAACCTATT 1594
Db      961  gacgaaggaatattgaacacacatcaatgaacacctctcaattttagaaccttgacgaactt 1020
OY      1595  GCTGCTATATGCGGTAAATTTTCAATGCTGTATGAGACAGGCTTTTCAACCCGCTGACA 1654
Db      1021  gctggtcatatcggttaataatcattcgaatgctgtatgacgaagctcttctaaacgtctgaca 1080
OY      1655  CGACTTGTAGCAATGCGCTTACGCTGCTGCCAAAGTTTACCGCTCTATGATGACCTTAC 1714
Db      1081  ccaacttgtagcaatgagccttaacgctgcgcgccaagtttaacgctctatgatgacttac 1140
OY      1715  CTGCTTGAAGGAAATCCAGAAAGATTGGAATTTACCAAGAGGGGCTCTACAGAAATGGA 1774
Db      1141  ctgctctgaaggaagcaatccagaggaattggaataatcaccagagggctctacagaaatgga 1200
OY      1775  GAATTCGTAATTTAAAAATGCTGCTGTTCTACGTCGGAACCTGCTCCACACAACAACCC 1834
Db      1201  gaattcgtatttaaaaaatggtgctcgttctacgttggaactcactgctccacaacaacccc 1260

```

```

RESULT 11
AAZ35939
ID      AAZ35939 standard; DNA; 1260 BP.
XX
AC      AAZ35939;
XX
XX      07-FEB-2000 (first entry)
XX      Streptococcus pneumoniae pbp1A TER isolate a) nucleotide sequence.
XX
XX      Streptococcus pneumoniae; penicillin binding protein; pbp2b; pbp1A;
XX      transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
XX      detection; identification; pneumococcal meningitis; ss.
OS      Streptococcus pneumoniae.
XX
XX      key Location/Qualifiers
FH

```

```

FT      CDS      1..1260
FT      FT      /*tag= a
FT      FT      /product= "pbp1A TER"
FT      FT      /note= "no stop codon given"
XX
XX      ZAG807024-A.
XX
XX      28-APR-1999.
XX
XX      05-AUG-1998; 98ZA-0007024.
XX
XX      01-AUG-1997; 97ZA-0006886.
XX
XX      (SAME-) SOUTH AFRICAN INST MEDICAL RES.
XX      (UYWI-) UNIV WITWATERSRAND.
XX      (MEDT-) MEDICAL RES COUNCIL.
XX
XX      Klugman KP, Smith AM, Du Plessis M;
XX
XX      WPI; 1999-601770/51.
XX      P-PSDB; AAY56106.
XX
XX      Polymerase chain reaction assays for detecting Streptococcus pneumonia
XX      useful for the diagnosis of pneumococcal meningitis
XX
XX      Claim 11, Fig 4; 63pp; English.
XX
XX      A polymerase chain reaction (PCR) assays have been developed for
XX      detecting an antibiotic resistant strain of Streptococcus pneumoniae
XX      using primers based on the penicillin binding protein 2b (pbp2b) gene
XX      and the pbp1A gene. The products and methods can be used for detecting
XX      S. pneumoniae, particularly antibiotic-resistant strains. They can be
XX      used for simultaneously diagnosing pneumococcal meningitis and
XX      identifying any antibiotic-resistant S. pneumoniae strains in a sample.
XX      The methods can be used for detecting S. pneumoniae strains resistant
XX      to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
XX      The assays can be adapted to detect other pathogens causing meningitis.
XX      The assays can be used to detect an antibiotic resistant strain of
XX      S. pneumoniae with a minimum inhibitory concentration (MIC) of
XX      0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
XX      a 224 bp product. The present sequence represents a Streptococcus
XX      pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
XX      sequence from the present invention.
XX
XX      Sequence 1260 BP; 405 A; 294 C; 261 G; 300 T; 0 other;
XX
XX      Query Match      41.9%; Score 838; DB 20; Length 1260;
XX      Best Local Similarity 99.4%; Pred. No. 0;
XX      Matches 1188; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

OY      617  CTACAAAGTCTCAANTAGCAAGTAATTTACCTGCTTACATGATGATTAATTTACCTCAAGAA 676
Db      43  ctacaagaagctcaaatcagcaagaatgaattacccgtgcttaacatgataattacctaagaagaa 102
OY      677  GTCAATCAATCAAGTTGAGAGAAAGAAAGCGCTATTAACCTACTCACACTGGAGATGATGC 736
Db      103  gtcatcaatcaagttgagaagaagaacagcgtataactaccactcaactgagatgagtcg 162
OY      737  TACCAAAATGTAGACCAAGAAAGCTCAAAACATCTGTGGGATATTTCATATACAGAGAA 796
Db      163  tacacaatgtlagaccagaagaagctcaaaaacatctgtyggatattacaaatagagaagaa 222
OY      797  TACGTCCTATATCAGACGATGAATTTCAAGTCGCTTCTACCATTTGTGATGTTTCTAAC 856
Db      223  taagtgtccatccagacgagatgaattgcaagtgctctaccatgttgatggtttcacaac 282
OY      857  GGTAAAGTCAATGGCCAGCTAGAGACAGCCATCAGTCAAGTAATGTTCTCTTGGAAAT 916
Db      283  ggtaaagtcaattgcccagctagagagcagcatcagtcagtaagtgttctctcggat 342
OY      917  AACCAACAGTAGAACAAACCGGACTGGGAGTCAACTGTAAACGATCACAGACTAT 976

```



```

Db 343 aaccaaagcagtagaanaaacccgcgactgaggatcaactatgaacacgaatcacagactat 402
Oy 977 GCTCCTGCTTGGAGTACGGTGTCTAGATTCACTGACTACTGCTTACGATGAGTGGCC 1036
Db 403 gctccgctcttgagtagcgtgtctcagtagcaactgcactcgtctcaagatgagccc 462
Oy 1037 TATTAACCTCCGGGAGCAAAATACCTCTGTTTATTAACCTGGGATAGGGGCTACTTTGGCAAC 1096
Db 463 tataactaccctggagaataataccctgtttataactggatagggtcactttgccaac 522
Oy 1097 ATCACTTTCGATACGCCCTCGACACAATGGCGAAAGTCCGACCGCTGGAAACTCTAAC 1156
Db 523 atcaacttgcataacgccttgcacaacatcgcaaacgtlccacgcgttggaaactcttaac 582
Oy 1157 AAGGTGGAAGTCAACCGCGCAAGACTTTCCTAAATGCTGTGAATTCGATCCCAAGT 1216
Db 583 aaggtcgagactcaacgcgcgaagacttccctaataatgctcggaaactgactacccaagt 642
Oy 1217 ATTCACTACTCAAAATGCCATTTCAAGTAAACACAAACCGAATCAGACAAAAAATATGAGAGA 1276
Db 643 attcaactcaaatgcatcattcaagtaacacacgcgaatcagacaaaaataatgagaca 702
Oy 1277 AGTAGGAAAAAGATGGCTCTGCTTACGCTGCTTGGCAATGGTGGAACTTACTATTA 1336
Db 703 agtagtaaaaagatggcttgcgtcttaccgcttgccttgcacaatggttgaacttactataa 762
Oy 1337 CCAATGTATTCATTAAGTTCGCTTACTAGTGGAGGAAAGAGTTCCTATATGTC 1396
Db 763 ccaagtataatccataaagctcgtctttagtagaggaagaaagatccatccatgctc 822
Oy 1397 GGAAGTCTGTCGACGAGAAAGACAGACGCTATATGATGACGACATGATGAAAAACAGTC 1456
Db 823 ggaactcgtgcataagaaagaaacagacgctatatatgacgcacatgataaagaaacagtc 882
Oy 1457 TTAGCTATGAGAACTGGAGAAATGCTATCTTGTGCTGGCTCCCTAGAGCTGTATAA 1516
Db 883 ttgagtagtaagaaactgagcaaaatgcttacttgccttgcctcccaaggctggtlaaaca 942
Oy 1517 GGAACCTCTAAGTATACAGACGAGAAATGAAACACATCAAGACCTCTCAATTTGTA 1576
Db 943 ggaacctcaactataacagagaggaatltgaaacacatacgaagccctcaatttga 1002
Oy 1577 GCACCTGATGAACATTTTCTGCTGCTATACGCGTAATATTCAAATGGCTGTATGACAGGC 1636
Db 1003 gaacctgaaactatltgctatacagctaaataltcaatgactgctatgagcaagc 1062
Oy 1637 TATTCCTAACGCTGTGACACACTTGTAGGCAATGGCTTACGCTGCTCCCAAGCTTAC 1696
Db 1063 tatctaaacgcttgcacacttltagcaatlgccttaccgctgcgtccaaagtttac 1122
Oy 1697 CCCTATGATGACCTACCTGCTGTGAAGAGCAATCCAGAGATTGGAATATACAGAG 1756
Db 1123 cgtctcatatgatacctactcgtctgaagaagaacatccagaagatltgatatataccagag 1182
Oy 1757 GGCGCTTCAGAAATGAGAAATTCGTATTAAAAATGGTCTGTTCTACGTGA 1811
Db 1183 gggctctacagaatgaggaattcgtatttaaaaatgctgcctgtcttaccgtgga 1237

```

RESULT 12
AAZ35947
ID AAZ35947 standard; DNA; 1260 BP.

AAZ35947;

07-FEB-2000 (first entry)

Streptococcus pneumoniae pDp1A TER isolate 1) nucleotide sequence.

Streptococcus pneumoniae: penicillin binding protein: pbp2B; pbp1A; transpeptidase encoding region; TER; antibiotic resistance; diagnosis; detection; identification; pneumococcal meningitis; ss.

XX

OS Streptococcus pneumoniae.

PN ZA9807024-A.

PD 28-APR-1999.

PF 05-AUG-1998; 982A-0007024.

PR 01-AUG-1997; 972A-0006886.

PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.

PA (UWI-) UNIV WITWATERSRAND.

PI (MEI-) MEDICAL RES COUNCIL.

PI Klugman KP, Smith AM, Du Plessis M;

DR WPI; 1999-601770/51.

PT Polymerase chain reaction assays for detecting Streptococcus pneumonia

PS Claim 11; Fig 4; 63pp; English.

A polymerase chain reaction (PCR) assays have been developed for detecting an antibiotic resistant strain of Streptococcus pneumoniae using primers based on the penicillin binding protein 2B (pbp2B) gene and the pDp1A gene. The products and methods can be used for detecting S. pneumoniae, particularly antibiotic-resistant strains. They can be used for simultaneously diagnosing pneumococcal meningitis and identifying any antibiotic-resistant S. pneumoniae strains in a sample. The methods can be used for detecting S. pneumoniae strains resistant to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin. The assays can be adapted to detect other pathogens causing meningitis. The assays can be used to detect an antibiotic resistant strain of S. pneumoniae with a minimum inhibitory concentration (MIC) of 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and a 224 bp product. The present sequence represents a Streptococcus pneumoniae pDp1A transpeptidase encoding region (TER) isolate nucleotide sequence from the present invention.

Sequence 1260 BP; 408 A; 286 C; 264 G; 302 T; 0 other;

Query Match 41.2%; Score 824; DB 20; Length 1260;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 824; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 617 CTACAAAGTCTCAAAATCAGCAAGTAATTAACCTGCTTACATGGATTAATTAAGTCAAGAA 676
Db 43 ctacaagctcacaatcagaagaatlaacctgcttaccatgataatlaacctcaagaa 102
Oy 677 GTCATTAATCAAGTTGAAGAAAGAAACAGGCTAATACCTCTGACACTGGAGTATGTC 736
Db 103 gtcatcaatcaagttgaagaagaacaggtctaaacctcaacacgctggaagatgctc 162
Oy 737 TACACAAATGTAGACCAAAAGCTCAAAAACATCTGTGGATTTTATACATACAGAGAA 796
Db 163 tacacaatgttagaccacaagaagctcaaaaacatctgtggatattacataaagcgaa 222
Oy 797 TACGTTGCTATTCAGACGATGATTAATGCAAGTGCCTTACCAATTTGATTTTAAAC 856
Db 223 tacgttgcctatccagaagatgaattgcaagtgcgtcttaccatgttgatgttctaac 282
Oy 857 GGTAAAGTATGTCCTCCAGCTAGAGACGCCATCATCAAGTAATTTTCCCTTGGAAAT 916
Db 283 ggtaaagtatgtccacagctagagcagccatcagtaagtaatgttcccttggaaat 342
Oy 917 AACCAAGCATAGTAAACCAACCGCAGCTGGGATGCAACATGATGAACCAAGTACAGACTAT 976
Db 343 aaccaaagcagtagaanaaacccgcgactgaggatcaactatgaacacgaatcacagact 402
Oy 977 GCTCCTGCTTGGAGTACGGTGTCTAGATTCACTGACTACTGCTTACGATGAGTGGCC 1036

```

Db 403 gctccgtcctgagtagcaggtgtctacagatcactcgtctacagatgagccc 462
QY 1037 TATACTACCCCTGGGCAAACTACTCCGTTTATACGGGATGGGGCTACTTGGCAAC 1096
Db 463 tataactaacctcggagaaatactccgttataaactggagtagggctacttggcaac 522
QY 1097 ATCACTTGGCAATACGCCCTCGCAACAATCGCAAACTCCCAACCCGCGGAAGCTTAAC 1156
Db 523 ataaccttgtaataacgcccgcgaacaatcgcaaaagctccagccggtggaaactctaac 582
QY 1157 AAGGTGGAGCTCAACCGCGCAAGACTTTCTTAATGCTGTAGGAATCGACTACCAAGT 1216
Db 583 aaggtcgaactcaaccgcgcgaagacttctctaattgtctagaatcgcactaccgaagt 642
QY 1217 ATTCACCTACCAATGATGCAATTAAGTAACACACGCAATCAGCAAAATAATGAGCA 1276
Db 643 atcactactcaaatgcattcaagtaatacaacacgaaatcgaacaataatgagca 702
QY 1277 AGTAGTGAAGAAGATGGCTGCTTACGCTTGGCAATGGTGGAACTTACTATATA 1336
Db 703 agtagtgaagaagatgctgtctgcttaagctgcttgcgaatggtggaacttactataa 762
QY 1337 CCATGTATATCCATAAAGTCGCTTTAGTAGGAGAGTGAAGAAAAGATTCTTAATGTC 1396
Db 763 ccaatgtataatccataaagtcgtcttaagtgatgagtagtgaagaaggtctctaatgct 822
QY 1397 GGAAGTCGTGCATGAAGAAAGACAGCCTATATGATGACCGA 1440
Db 823 ggaactcgtgcatagaagaagaacagcctataatgatgacga 866

RESULT 13
AA083241
ID AA083241 standard; DNA; 960 BP.
XX
XX AA083241:
DT 14-FEB-1996 (first entry)
XX
XX Streptococcus pneumoniae strain SPRU42 Exp2 DNA.
DE
XX
XX Exp2: export protein; permease like protein;
KW penicillin binding protein 1A; pbp1a; pbp1; exp1; exp3; pad1;
KW pyruvate oxidase; virulence determinant; regulatory element;
KW acellular vaccine; antibody; ds.
XX
XX OS Streptococcus pneumoniae.
XX
XX PN W09506732-A2.
XX
XX PD 09-MAR-1995.
XX
XX PF 01-SEP-1994; 94WO-US09942.
XX
XX PR 01-SEP-1993; 93US-0116541.
XX 18-MAY-1994; 94US-0245511.
XX PA (UYRO) UNIV ROCKEFELLER.
XX
XX PI Masure HR, Pearce BJ, Tuomanen E;
XX
XX DR WPI; 1995-115448/15.
XX P-PSDB; AAR70153.
XX
XX PT Novel gene fragments encoding specific bacterial exported proteins
XX - specifically of S. pneumoniae; useful as vaccines
XX
XX PS Claim 28; Page 88-9; 168pp; English.
XX
XX CC This sequence represents the DNA encoding exp2. This sequence is
XX identical to that for ponA which encodes penicillin-binding protein 1A
XX (pbp1a). The protein encoded by this sequence is involved in adhesion
XX of bacteria to target cells. This sequence encodes an exported protein

CC of S.pneumoniae. Export proteins are the proteins in pathogenic
CC bacteria that are virulence determinants. Other export proteins include
CC pbpA (see AAR70152), exp1, exp3, and pad1 (encoded by the sequence shown
CC in AA083259). This sequence can be inserted into an expression vector
CC (preferably a bacterial expression vector) to provide for high levels of
CC expression of the protein. The protein can then be used in the
CC production of an acellular vaccine. These vaccines are used to provide
CC protection from Gram positive bacterial infection. Antibodies against
CC export proteins can be used for diagnosis of infection and in passive
CC immune therapy.
XX
XX SQ Sequence 960 BP; 315 A; 224 C; 182 G; 239 T; 0 other:
Query Match 39.8%; Score 796; DB 16; Length 960;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TAAATCTACGACAAATTAATCACTCAATGCTGACTTGGTTCGAAACCGCGCTCAA 60
Db 12 taaatctacgacaataaataatcaactcattgctgacttggcttgaagccgcgtca 71
QY 61 TGCCCAAGCTAATGATATTTCCACAGATTGGTTAAGCAATGCTTTCTATCGAAGCA 120
Db 72 tgcccaagctaatgatatctccacagatttggtaaagcattgttctatcgaaagca 131
QY 121 TCGCTTCTTGACACACAGGGGGATTGATACATCCGATCCGTGGAGCTTCTTGCGCA 180
Db 132 tgcgttcttgacacacaggggattgataccatccgatatccgtggagcttcttggca 191
QY 181 TGTGCAAGCAATTCCTCCAGGTGATCACTTCACCAACAGTTGATTAAGTTGAC 240
Db 192 tctgcaagcaattccctccaaagtgtgactcagctcctcaacaaagttgattgac 251
QY 241 TTACTTTTCACTTCGACCTCCGACAGACTATTTCGTAAGGCTCAGGAAGCTTGCT 300
Db 252 ttactttcaactcgacttcgcgaagactatttccgtaaggtctaaagagcttggct 311
QY 301 AGCGATTCAGTTAGAACAAACCAACCAAGCAAGAAATCTTGACCTACTATTAATAA 360
Db 312 agcgattcagttagaacaaagaagaacaaagaatcttgcctcactataataa 371
QY 361 GGTCTACATGTCGTAATGGAAGTATGGAATGCAGACAGCTCAAAATCTATGAGTAA 420
Db 372 ggtctacatgcttaatggaagctatgaaatgcagacagcagcctcaaaactatgtaa 431
QY 421 AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTGTGCTGCTGGAATGGCTCAGGACC 480
Db 432 agacctcaatatttaagtttacctcagttagcttgcgtgcgaaatgctcgaagacc 491
QY 481 AAACCAATATGACCCGCTATTCACATCCAGAGGACGCCAGCAAGCTTGAGTGTGCTT 540
Db 492 aaaccaatatagaccctattcacatccagaagaagcccaagaccgcgaacttggctc 551
QY 541 ATCTGAAATGAAAAATCAAGGCTACATCTGCTGGAACAGTATGAGAAAGCACTCAATAC 600
Db 552 atctgaaatgaaaaatcaaggtctacatctctgctgacagtagtagaagaagcagctaac 611
QY 601 ACCAATTACTGATGAGCTACAAAGTCTCAATATGACGAAGTAAATTAACCTCTACATGGA 660
Db 612 accaattactgatgagctacaaagctcaaatcagaagtaattaccctgcttaccgga 671
QY 661 TAAATTAAGTGAAGTCAATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTAC 720
Db 672 taattactgaagaaggtcactcaatcaagttgagaagaagaacgctactactaccac 731
QY 721 AACTGGATGATGTCTACCAAAATGTAGACGAAGAAGCTCAAAAACATCTGTGGATAT 780
Db 732 aactggatgatgtctacaaaatgtagacgaagaagctcaaaaacatctgtggatatt 791
QY 781 TTACAATACAGAGATAGCTTGCCTATCCAGACGATGAATTTGCAAGTCCGCTTACCAT 840
Db 792 ttacaatacagagatagcttgcctatccagacgataagttcaagtcgcttaccat 851

QY 841 TGTGATGTTTCTAACGTTAAAGTCAATTCGCGCTAGAGACGACCATCAAGTAA 900
|||||
DB 852 tgttgatgtttcttaacgagtaagtcattgcccagctagagacgcacatcagtaagtaa 911
|||||
QY 901 TGTTCCTTCGGAATTACCAACGACGATGAAACAAACCGGACTGGGGA 949
|||||
DB 912 tgttccttcggaattacacagcagtagaacaacgcgactggyga 960
|||||
RESULT 14
AAZ35948
ID AAZ35948 standard; DNA; 1260 BP.
XX
AC AAZ35948;
XX
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1A TER isolate j) nucleotide sequence.
XX
KM Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
KM transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KM detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UWI-) UNIT WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX
DR WPI; 1999-601770/51.
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX
XX
PS Claim 11: Fig 4; 63pp; English.
XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 403 A; 285 C; 264 G; 308 T; 0 other;

Query Match 38.2%; Score 764; DB 20; Length 1260;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 864; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 575 GAACAGTATGAGAAAGACGATCAATACCAATTTGATGAGACTACAAAGTCTCAATCA 634
|||||

DB 1 gaacaglatgagaagcagclcaatacaccaattactgtagacatacaaatgctcaaatca 60
QY 635 GCAAGTAAATTAACCTGCTTAACATGATGATTAATTAACCTGAGAAATCAATCAAGTTGA 694
|||||
DB 61 gcaagtaataccctgcttaacatgataatlaacccaaggaatcaatcaatgaa 120
|||||
QY 695 GAAGAACAAGCTTAATTAACCTACTACAACTGGATGGATGTCACAAATGTAGACCA 754
|||||
DB 121 gaagaaacagcgtataacccgtcacacatggaatgattgctcaacaatgtagacca 180
|||||
QY 755 GAAGCTCAAAAACATCTGTGGGATATTTCATATACAGACGATAGCTTGCTTCCAGAC 814
|||||
DB 181 gaagctcaaaaacatctgtggaatttaacaacagacgaatagctgtgcttaaccaag 240
|||||
QY 815 GATGAATTCGAATTCGCTTCTTACCATTTGATGATTTTGAAGCTTAAGCTTAAGCTTGGCCAG 874
|||||
DB 241 gatgaattgcaatgcttcttaccattgttgaatgcttcaagcttaagctatgcccag 300
|||||
QY 875 CTAGAGACAGCCATCAGTCAAGTAAATGTTCTTCGGAATTAAACAGCAGTGAACA 934
|||||
DB 301 ctagagacagccatcagtaagtaatgcttcttgcgaattacaagcagtagaagaaca 360
|||||
QY 935 AACCGGACTGGGATCAACTATGAACCGATACAGACTATGCTGCTGCTTGAGATAC 994
|||||
DB 361 aaccggaactggygatcaactatgaaacgatacagactatgctcctgcttgagatc 420
|||||
QY 995 GGTGCTAGATTCATCACTGCTACTATCGTTACAGATGAGCCCTATTAATCTCCCTGGGACA 1054
|||||
DB 421 ggtgctagatTCATCACTGCTACTATCGTTACAGATGAGCCCTATTAATCTCCCTGGGACA 480
|||||
QY 1055 AATACTCTGTTTAACTAGTGGATAGGGGCTACTTTGGCAACATCACTTGCAATACGCC 1114
|||||
DB 481 aatactctgttttaactatcagtagagtagggtactttgcaacaatcacttgcaatagcc 540
|||||
QY 1115 CTGCAACAATCGGGAACAGTCCAGCCGTTGGAACACTTAACAAGTGGCAGCTCAACGCC 1174
|||||
DB 541 ctgcaacaatcgggaacagtcaccagcgtggaaccttaacaaggctcggaactcaacgcg 600
|||||
QY 1175 GCCAAGACTTCTTAATGCTTAGAATGAGATGATACCAAGTATTCATCTCAATATGCC 1234
|||||
DB 601 gccaaagacttcttaaatgcttagaattgagatcgaatcgaactcaactcaactcaatgccc 660
|||||
QY 1235 AATTCAAGTTAACCAACCGATCAGACAAAATATGAGCAAGTATGTAAGATGCT 1294
|||||
DB 661 attcaagtaacacacacggaatcagaacaaaataatgagcgaatgtagaagaagcgt 720
|||||
QY 1295 GCTGCTTACGCTGCTTTGCAATGCTGGAATCTTACTATTAACCAATGTATATCAATAA 1354
|||||
DB 721 gctgcttacgctgcttgcgaatgtagaactactataaccaatgatatcatcaaa 780
|||||
QY 1355 GTGCTTTTATGATGAGAGTGAAGAAAGATTCTTAATGTGGAACCTGTCGCAATGAAG 1414
|||||
DB 781 gtgcttttagatgagtaggagtagaagaagttctcaatgctcggaactcgtgcacgaag 840
|||||
QY 1415 GAACAGACAGCCATATGATGATGACCA 1440
|||||
DB 841 gaacgacagcgtatgatgaccca 866
|||||
RESULT 15
AAZ35949
ID AAZ35949 standard; DNA; 1260 BP.
XX
AC AAZ35949;
XX
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1A TER isolate k) nucleotide sequence.
XX
KM Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
KM transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KM detection; identification; pneumococcal meningitis; ss.
XX

Query Match	Similarity	37.3%	Score 745	DB 20	Length 1260
Best Local	Similarity	99.8%	Pred. No. 0		
Matches	845	Conservative	0	Mismatches	2
				IndeIs	0
				Gaps	0
OY	594	TCGAATACACCAATTTACTGATGCATACCAAGCTCTAAATCAGACAAGTAATTAACCTGGCTT	653		
Db	20	tcaatacaccaattactgtgtgactacaaagctccaatccagcaagtaattacccttgctt	79		
OY	654	ACATGGATTAAATTACCTCAAGAGATCTAATCAAGTTGAAGAGAAACAGGCTATTACC	713		
Db	80	acatgataattaccctccaaggaagatcatcaatcaaglttgaagaagaacagctataacc	139		
OY	714	TACTACACAACCTGGATGATGTCTACACAAATGTATGACCAGAAGCTCAAAAACATCTGT	773		
Db	140	tgtctacaacactggatggtgtgtctctacacaatgtagaccagaagctccaataacatctgt	199		
OY	774	GGGATTATTACATATACAGAGCAATACGTTGCTATCCACAGATGAATTGCAGATCGCTT	833		
Db	200	gggataatttacaatacagaagcatcagttgctctatccagaagatgaattgcaagctgcct	259		
OY	834	CTACCATTTGTTATGTTTCTTAACGGCTAAAGTATTGGCCAGGTAGAGACGCCATTCAGT	893		
Db	260	ctaccatcgttgatgtttcttcaacggttaaacgtatcgtccagctagaagacggcatcagct	319		
OY	894	CAAGTAATGTTTCCCTTCGGAATTAACCAAGCAGTAGAAGAAACAACCGCGACTGGGATCAA	953		
Db	320	caagtaatgtttccctcttggaattaaaccaagcgtatgaatacaaacggcgactgtggatcaa	379		
OY	954	CTATGAAAACGATCACAGACTATGCTCCTGCTTGAGTAGGCTGTTACGATTCACATCG	1013		

XX	AAH02060	RESULT 16
ID	AAH02060 standard; DNA; 930 BP.	
XX		
AC	AAH02060;	
XX		
DT	24-JUL-2001 (first entry)	
XX		
DE	Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2053.	
XX		
KW	Species specific; genus specific; family specific; probe; detection;	
KW	identification; algal; archaeal; bacterial; fungal; parasitical;	
KW	microorganism; diagnosis; translation elongation factor Tu; toxin;	
KW	translational elongation factor G; RecA recombinase; resistance;	
KW	catalytic subunit of proton-translocating ATPase; antimicrobial;	
KW	vaccine; primer; ds.	
XX		
OS	Streptococcus pneumoniae.	
XX		
PN	W0200123604-A2.	
XX		
PD	05-APR-2001.	
XX		
PF	28-SEP-2000; 2000MO-CA01150.	
XX		
PR	28-SEP-1999; 99CA-2283458.	
PR	19-MAY-2000; 2000CA-2307010.	
XX		
PA	(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.	
XX		
PI	Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;	
PI	Picard FJ, Roy PH.	
XX		
DR	WPI; 2001-245006/25.	
XX		
PT	Nucleic acid sequences are used to generate universal probes and	
PT	primers which can be used to identify and detect the presence of algal	
PT	archaeal, bacterial, fungal and parasitcal species in a test sample -	

XX Disclosure; Page 1468; 1580pp; English.
XX
XX The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atp and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitical
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitical species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp. using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
XX
SQ Sequence 930 BP; 289 A; 212 C; 196 G; 233 T; 0 other;

Query Match 31.0%; Score 620; DB 22; Length 930;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 770 CTTGGGATATTTCATATACAGCAATACGTTGCTATCCAGCATGATGATGCAATC 829
DB 1 CTTGGGATATTTCATATACAGCAATACGTTGCTATCCAGCATGATGATGCAATC 829
QY 830 GCTTACCATGTTGATGTTCTACAGGTAAGATGCTAGGAGTACAGCAGCCAT 889
DB 61 GCTTACCATGTTGATGTTCTACAGGTAAGATGCTAGGAGTACAGCAGCCAT 889
QY 890 CAGTCAAGTAAATGTTTCTTGGGAATTAACAGAGTAAACAAACCGCATGGGGA 949
DB 121 CAGTCAAGTAAATGTTTCTTGGGAATTAACAGAGTAAACAAACCGCATGGGGA 949
QY 950 TCACATATGAACCGATTCACAGACTGTGCTGCTGGAGTACGTTGATGATTTCA 1009
DB 181 TCACATATGAACCGATTCACAGACTGTGCTGCTGGAGTACGTTGATGATTTCA 1009
QY 1010 ACTGCTACTATGCTTACGATGAGCCCTTAATACCTCGGGAATTAATCTCGTTTAT 1069
DB 241 ACTGCTACTATGCTTACGATGAGCCCTTAATACCTCGGGAATTAATCTCGTTTAT 1069
QY 1070 AACTGGGATAGGGGCTACTTTGGCAACATACCTTGGCAATAGCCCTGGAACATGCGGA 1129
DB 301 AACTGGGATAGGGGCTACTTTGGCAACATACCTTGGCAATAGCCCTGGAACATGCGGA 1129
QY 1130 AACGTCGCCGCGGTGGAACCTTAACAGAGTGGGCTAACCGGCCCAAGACTTTCCTTA 1189
DB 361 AACGTCGCCGCGGTGGAACCTTAACAGAGTGGGCTAACCGGCCCAAGACTTTCCTTA 1189
QY 1190 AATGCTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249
DB 421 AATGCTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249
QY 1250 ACCGAATGAGCAAAAAATATGAGCAAGTATGTAAGATGGCTGCTGATACGCTGCC 1309
DB 481 ACCGAATGAGCAAAAAATATGAGCAAGTATGTAAGATGGCTGCTGATACGCTGCC 1309
QY 1310 TTTCGCAATGTTGAGACTTACTATTAACCAATGATATATCCATTAAGTCTGTTTACTGAT 1369

DB 541 TTTCGCAATGTTGAGACTTACTATTAACCAATGATATATCCATTAAGTCTGTTTACTGAT 1369
QY 1370 GGGAGTGAAGAAAGATGCTTATGTCGGAACCTGTCGATGATGATGATGATGATGATGAT 1429
DB 601 GGGAGTGAAGAAAGATGCTTATGTCGGAACCTGTCGATGATGATGATGATGATGATGAT 1429
QY 1430 ATGATGACCGA 1440
DB 661 ATGATGACCGA 671

RESULT 17
AAH01023
ID AAH01023 standard; DNA; 1199 BP.
XX
XX AAH01023;
AC
XX 24-JUL-2001 (first entry)
XX
XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1014.
DE
XX
XX Species specific; genus specific; family specific; probe; detection;
KW Identification; algal; archaeal; bacterial; fungal; parasitical;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
XX Streptococcus pneumoniae.
XX
XX WO200123604-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-CA01150.
XX
XX 28-SEP-1999; 99CA-2283458.
XX
XX 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
XX WPI: 2001-245006/25.
PS
XX
XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitical species in a test sample -
XX
XX
XX Claim 27: Page 968; 1580pp; English.
PS
XX
XX The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atp and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitical
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitical species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,

CC *Neisseria gonorrhoeae* and *Staphylococcus* sp., using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.

SQ Sequence 1199 BP; 386 A; 286 C; 240 G; 287 T; 0 other;

Query Match	29.5%;	Score 590;	DB 22;	Length 1199;
Best Local Similarity	99.7%;	Pred. No. 4.1e-294;		
Matches 690;	Conservative	0;	Mismatches 2;	Indels 0;
			Gaps	0;

D	b	1	caaaaacatctgtyggagattactacaatacagacgaatacgttgcctatccagagatga	60
Q	y	761	CAAAAACATCTGGTGGGATATTATTCAATTAAGACAGAAATACGTTGACCTATCCAAACATGAA	820
D	b	1	caaaaacatctgtyggagattactacaatacagacgaatacgttgcctatccagagatga	60
Q	y	821	TTGCAAGTCGCTTCACACTATTGTTGATGTTTCTAACGGTAAGTCAATTGCCAGCTAGCA	880
D	b	61	ttgcaagtcgcttctacacattgttgttcttcaacgtaaaagtaatttgcgccagctaga	120
Q	y	881	GCACGCCATCAGTCAACTAATGTTTCCCTTCGGAAATTAACCAAGCAAGTGAAMAACAACGC	940
D	b	121	gcacgccatcagtcgaagtaattgttccctcggaaataaccgaagcagtagaacaacacgc	180
Q	y	941	GACTGGGGATCAACTATGAAACCGATCACAGACTATGTCCTGCTTGGAGTACGGTTC	1000
D	b	181	gactgggagatcaactatgaaaccgatcacagactatgctctcgtccttggatgagtytc	240
Q	y	1001	TACATTCACACTGCTACTATACGCTTCACAGATGAGGCCCTAATACCTCCGGACAAATACT	1060
D	b	241	tacagattcaactgtctactacatcgttcaagatgagcccttaactaccctggagcaataacc	300
Q	y	1061	CTGTATTATTAACGTGGATAGGGGCTACTTGGGAAACATCACTTGCATATACGCCCTCGCA	1120
D	b	301	ccgtttataactcgtgagtagggactacttggacaatacactccttgcataacgccttgaa	360
Q	y	1121	CAATTCGGAAACGTCCAGCCGTGGAAACTCTAACAAAGTGGACTCAACCGCGCCAA	1180
D	b	361	caatcgcgaaacgltcccaagcgttggaaactctaaacaagltcggactcaacgcgcgcaag	420
Q	y	1181	ACTTTCCTAATGGCTTAGCAATTCGCAATCCCAAGATTCACCTACACAAATCCCATTTCA	1240
D	b	421	acttctctaattgtctcggaaatcgcgactaccaccaagatatacctaactcaaatgtccattca	480
Q	y	1241	AGTAACCAACAACCGAATCAGACAAAATAATGAGCAAGTAGTGAAAGAATGGCTGCTGCT	1300
D	b	481	agtaaccaacaacggaatcagaacaaaataatgagcaagtagtgaanaagatgcttgcct	540
Q	y	1301	TAGCGTCCTTGGCAATGGTGGAACTTACTTAACCAAACTAATATTCATTAAGTCGTC	1360
D	b	541	taagtcgcttgcgaaatgtgtggaactactactaataacaaatgtatataccttaaaagtgc	600
Q	y	1361	TTTGTGTATGGGAGTGAATAAAGATTCCTTAATGTGTGGAATCGTGCATGAGGAAGC	1420
D	b	601	tttagtgtgtagtgatgaaaagagttctctaatgtcggaaatcgtgcgcatgagaagaaacg	660
Q	y	1421	ACAGCCTATATGATGACCGCATGATGAAAC 1452	
D	b	661	acagcctatatgtatgaccgcgatgagaanaac 692	

RESULT	18
AA235945	
ID	AA235945 standard; DNA; 1260 BP.
XX	
AC	AA235945;
XX	
DT	
XX	
DE	07-FEB-2000 (first entry)
XX	
XW	Streptococcus pneumoniae pbp1A TER isolate g) nucleotide sequence.
XX	
TW	Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A

KM transpeptidase encoding region; TEK; antibiotic resistance; diagnosis;
KM detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.

Streptococcus pneumoniae.

PN ZA9807024-A

PD 28-APR-1999

PF 05-AUG-1998; 98ZA-0007024.

PR 01-AUG-1997; 97ZA-0006886.

PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.

PA (MEDI-) MEDICAL RES COUNCIL.

PA (MEDI-) MEDICAL RES COUNCIL.

Klugman KP, Smith AM, Du Plessis M,

DR WPT; 1999-601770/51.

PT Polymerase chain reaction assays for detecting *Streptococcus pneumoniae*

PS Claim 11; Fig 4; 63pp; English.

CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of *Streptococcus pneumoniae*
CC using primers based on the penicillin binding protein 2b (pbp2B) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC *S. pneumoniae*, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant *S. pneumoniae* strains in a sample.
CC The methods can be used for detecting *S. pneumoniae* strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC *S. pneumoniae* with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a *Streptococcus*
CC *pneumoniae* pbp1a transpeptidase encoding region (TTE) isolate nucleotide
CC sequence from the present invention.

SQ Sequence 1260 BP; 404 A; 281 C; 268 G; 307 T; 0 other;

Query Match	29.1%;	Score 581;	DB 20;	length 1260;
Best Local Similarity	99.7%;	Pred. No. 1.8e-289;		
Matches 681; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0

Db	166	acaaatgtagaccagaagctccaataacatctgttggaattattacaatcagagaatc	225
Qy	740	ACAATGTAGACCCAGAAAGCTCAAAAAATCTGGATTATTACATACAGCAATAC	799
Db	166	acaaatgtagaccagaagctccaataacatctgttggaattattacaatcagagaatc	225
Qy	800	GTTGCGCTATCCAGACGATGAATTGGCAAGTCGCTTCTACCATTTGTTGATGTTCTTAACGCT	859
Db	226	gttcaccatccagacgaatgaaatctgaaagtcgaagtcctccaccattgttgaattcttaacggt	285
Qy	860	AAATCATTTGCCAGCTAGAGGACGCCCATCATCTCAAGTAATGTTTCCCTTCGGAATTAC	919
Db	286	aaagtactatggccagctcagtagaagaagccatcagtcgaatgaattgttcccttggaattaac	345
Qy	920	CAACGACTAGAAACAACCCGCACTGGGGATCACTATGTAAGAACCGATCAGACACTATGCT	979
Db	346	caagcaatagaacaacaaccctgcacttgggatacactatgaataaccgatacagaactatgct	405
Qy	980	CCCTGCTTGGAGTAGAGGCTGTCAGATTCAACGTACACTATGTTTACACATGAGCCCAT	1033
Db	406	ctgtcctcttgagatcagtgctctcagaattcaactgcactactatcglttcagaatgagcccat	465
Qy	1040	AACATCCCTGGACAATACTCCTGTTTATTAAGTGGATAGGGGCTACTTTGGCAATC	1099
Db	466	aactacccttgggacaataactctctgttcaactaacttggaatagggcactacttggcaacac	525

PI Black MT, Hodgson JE, Knowles DUC, Lonetto MA, Nicholas RO;
PI Stodola RK;
XX
XX
DR WPI: 1998-159452/14.
DR P-PSDB; AA85958, AA85959.
XX
XX
PT Streptococcus pneumoniae proteins and related DNA - useful for
PT screening compounds for antibacterial activity
XX
XX
PS Claim 4; Page 171-172; 640pp; English.
XX
XX
CC This invention describes novel isolated Streptococcus pneumoniae
CC polynucleotides (see AA96173-796494) and their encoded proteins (see
CC AA95792-Y86182). The DNA, vectors and host cells described in the
CC method of the invention are useful for the recombinant expression of the
CC polypeptides. The polypeptides are useful for treatment or prevention of
CC disease, or diagnosis of disease related to expression or activity of
CC such a polypeptide. They can also be used to screen for compounds which
CC interact with and inhibit or activate such a polypeptide. The
CC polypeptides (or DNA encoding them, via gene therapy) are also useful
CC for inducing an immunological response in a mammal. The antagonists are
CC useful to inhibit such bacterial polypeptides. The polypeptides are
CC particularly useful to identify antimicrobial compounds and antibiotics.
CC They are also useful to determine their role in pathogenesis of
CC infection, dysfunction and disease.
XX
XX
SQ Sequence 2172 BP; 629 A; 498 C; 409 G; 635 T; 1 other;

Query Match 27.1%; Score 541; DB 19; Length 2172;
Best Local Similarity 99.5%; Pred. No. 8,4e-269;
Matches 811; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 1186 CCAATATGGCTAGCAATGACATCCCAAGTATTCCTACTCAATTAATGCAATTCAGTAA 1245
DB 13 cctaaatgctcaggaatcgaactaccacgaatctcactcactcaatcgaagtaa 72
OY 1246 CACACCGCATGACACAAAAAATATGAGCAAGTATGAAAAGATGGCTGCTTACGC 1305
DB 73 caaacgcgacatcgacacaaaataatgagcaagtagtgaagaatgctgcgtctaacgc 132
OY 1306 TGGCTTGGCAATGGTGGCACTTACTATAACCAATGTATATCCATAAAGTGTCTTTAG 1365
DB 133 tgccttgcataatgctggaactataacacgaatataccataaagtcgtcttag 192
OY 1366 TGATGGAGTGAAGAAAGCTCTCTATATGCGAAGTCTGCAATGAAGAAAGCAGAC 1425
DB 193 tgaatggaatgaaagaagatctcctaagtcggaactcgcgaagaagaagcagcgc 252
OY 1426 CTATATGATGACCGACATGATGAAAAACAGTCTTGACTTATGAACTGAGAG- AATGCGT 1484
DB 253 catatgatgcgcgacatgatgaagaacagctcttgagttatggaacgcggaataagcct 312
OY 1485 ATCTGCTTGGCTCCTCGAGGCTGGTAAACAGAACTCTTACTATACAGAGAGGANA 1544
DB 313 atcttgcttggtcctcctcagctggtlaaaacggaacacctactaatacagaagagaa 372
OY 1545 TTGAAACACACATCAAGACCTCTCAATTTGTAGACCTGATGAACATTTGGTGGGTATA 1604
DB 373 ttgaaacacacatcaagacctcctcaattgttagcacctgagaaacatttgcgtgctata 432
OY 1605 CCGGTAATATTCATGAGTGTATGAGACAGGCTATTTCAACGCTGACACACATTTGAG 1664
DB 433 cgcgtaaatatcatcatatgctgtatgagacagctattctcaacgctcgaacacattgtg 492
OY 1665 GCAATGCTTACGCTGCTGCCAAAGTTTACCGCTATATGATGACCTTACCTGTGAAG 1724
DB 493 gaatgcttactagctgcgtgcgaagttaccgcctctatgtagcactaccctgcctgaag 552
OY 1725 GAAGCAATCCAGAAAGTTGGATATATCCAGAGGCTCTACAGAAATGAGAAATTCGTAT 1784
DB 553 gaagcaatccagagattggaatatataccagagggctctacagaatggaattcgtat 612

OY 1785 TTAATAATGATGCTGCTTCTAGCTGAGACTGACCTGCTCCACAGAACCCCATCACTG 1844
DB 613 ttaaaatgctgctgcttctagctggaactaacctgctcacaacaccccatcaacg 672
OY 1845 AAAGTTCAAGCTCATCATCATAGATAGTTCAACTTACAGCTGTAGCTACACCACTCAGCA 1904
DB 673 aaagtcaagctcatcatcatcatagatagttcaacttcaacgcttagctcaacacccaaga 732
OY 1905 CAAATATATGTAGTACGACTTACCAATCCCTAACATATATGCGACAAATATACAAACCCG 1964
DB 733 caaataatagctgactaccatccatacaataaatacgaacatacaatacaaccccg 792
OY 1965 ATCAACAAATGAGAAATTCCTCAACACGACACCAACA 1999
DB 793 atcaacaaatcgaatcctcctcaacgcagacacaacca 827

RESULT 22

AAH02054
AAH02054 standard; DNA; 2160 BP.

AAH02054;

24-JUL-2001 (first entry)

Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2047.

Species specific; genus specific; family specific; probe: detection;
Identification; algal; archaeal; bacterial; fungal; parasitical;
Microorganism; diagnosis; translation elongation factor Tu; toxin;
translational elongation factor G; RecA recombinase; resistance;
catalytic subunit of proton-translocating ATPase; antimicrobial;
vaccine; primer; ds.

Streptococcus pneumoniae.

WO200123604-A2.

05-APR-2001.

28-SEP-2000; 2000MO-CA01150.

28-SEP-1999; 99CA-2283458.

19-MAY-2000; 2000CA-2307010.

(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.

Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
Picard FJ, Roy PH;

WPI: 2001-245006/25.

Nucleic acid sequences are used to generate universal probes and
primers which can be used to identify and detect the presence of algal,
archaeal, bacterial, fungal and parasitological species in a test sample -
Disclosure; Page 1462-1463; 1580pp; English.

The present invention describes a method for generating a repository of
nucleic acids of tu, fus, atp and/or recA genes from which probes
and/or primers are derived. The method comprises amplifying the nucleic
acids of determined algal, archaeal, bacterial, fungal and parasitological
species with a combination of defined primer pairs. The method can be
used for producing probes and/or primers for detecting one or more
related microorganisms e.g. algae, archaea, bacteria, fungi and
parasites, for universal detection and for specific and ubiquitous
detection and identification of an algal, archaeal, bacterial, fungal
and parasitological species, genus, family and group. A nucleic acid (i)
obtained using the method of the invention can be used for the universal
detection of any bacterium, fungus or parasite in a sample and for the
detection of at least one antimicrobial agent resistance gene or at
least one toxin gene. hexa nucleic acids are used for the specific and
ubiquitous detection and for identification of Streptococcus pneumoniae.

CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
XX
SQ Sequence 2160 BP; 701 A; 494 C; 395 G; 570 T; 0 other;

Query Match 19.8%; Score 396; DB 22; Length 2160;
Best Local Similarity 99.1%; Pred. No. 6,3e-194;
Matches 696; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TAAATCTACGACAAATAAATCACTGCTGACTTGGGTTCTGAACGCCGCTCAA 60
DB 159 taaatctacgacataaaatcaatcattgctgacttggtctgaagcgcgctcaa 218
QY 61 TGCCCAAGCTAATATATTCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACA 120
DB 219 tgcccaagctaatatattccacagatttggtlaagcaatcgcttctatcgaagaca 278
QY 121 TCGCTTCTGACACAGAGGGGATTTGATACCATCGTATCCTGGAGCTTTCTGGCCAA 180
DB 279 tcgcttctgacacaggggattgatacatccgtatcccttgaggcttcttgccaa 338
QY 181 TCTGCAAGCAATTCCTCCCAAGTGATCATCTCATCCACAGATTGATTGATTGAC 240
DB 339 tctgcaagaatccctccaagtgatcaactcaccacaagatgataatctgac 398
QY 241 TTTACTTTCACTGCTCCGACGACGACTATTTCTCGTAAGGCTCAGGAAGTTGGTT 300
DB 399 ttactttcaactcgctcccgacagactatttcgtaaggtccagaagcttggtt 458
QY 301 AGCGATTCACTTGAACAAAAGCAACCAAGAAATCTTGACTATATATAATAA 360
DB 459 agcgattcagttgaacaaaagcaacaaagaaatcttgacctatataataa 518
QY 361 GGTTTACATGTTTATGGGAATTCGATGACAGACGAGCTAAAATCTATGTGTA 420
DB 519 ggttatacttcaatgacactatgaaagacagcagctcaaaactactatgttaa 578
QY 421 AGACCTCAATAATTTAGTTTACCTCAGTTAGCTTGCTGGTGAATGCTCAGGAC 480
DB 579 agacctcaataatlaagttactcagttagcttctgctgagaaatgctccagcac 638
QY 481 AAACCAATATGACCCCTATTTCATCCAGAGACGCCCAAGCCGGAACCTTGCTT 540
DB 639 aaaccaatattgacccctattcaatccagagaagcccaagacgcgaaacttgctt 698
QY 541 ATCTGAATGAAAAATTAAGCTATCTGCTGAACAGTATGAGAAGCAATCAATAC 600
DB 699 atctgaatgaaaaatcaagttactcctcgtcgtgacagatagagaagcaatcac 758
QY 601 ACCAATACGATGAGTACAAAAGTCAATCAATGCAAGTAATACCTGTTTCATGGA 660
DB 759 accaatactcgtatggtactaataagttcacaatcagcaagtaataacctgttaccatgga 818
QY 661 TAAATTACCTCAGAGATCATCAATCAAGTTGAAGAAGAAC 702
DB 819 taattacctcaaggaagtcataatcaagttgaagaagaac 860

RESULT 23
AAH02147
ID AAH02147 standard; DNA: 2157 BP.
AC
XX AAH02147;
XX

DT 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2140.
XX
KW Species specific; genus specific; family specific; probe; detection;
KW Identification; algal; archaeal; bacterial; fungal; parasitical;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
XX Streptococcus pneumoniae.
XX
XX WO200123604-A2.
XX
XX 05-Apr-2001.
XX
XX 28-SEP-2000; 2000WO-CA01150.
XX
XX 28-SEP-1999; 99CA-2283458.
XX
XX 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
XX Picard FJ, Roy PH;
XX WPL; 2001-245006/25.
XX
XX
XX Nucleic acid sequences are used to generate universal probes and
XX primers which can be used to identify and detect the presence of algal,
XX PT archaeal, bacterial, fungal and parasitcal species in a test sample -
XX PS Disclosure; Page 1508-1509; 1580pp; English.
XX
XX The present invention describes a method for generating a repertory of
XX nucleic acids of tuf, fus, atpd and/or recA genes from which probes
XX and/or primers are derived. The method comprises amplifying the nucleic
XX acids of determined algal, archaeal, bacterial, fungal and parasitcal
XX species with a combination of defined primer pairs. The method can be
XX used for producing probes and/or primers for detecting one or more
XX related microorganisms e.g. algae, archaea, bacteria, fungi and
XX parasites, for universal detection and for specific and ubiquitous
XX detection and identification of an algal, archaeal, bacterial, fungal
XX CC and parasitcal species, genus, family and group. A nucleic acid (1)
XX obtained using the method of the invention can be used for the universal
XX CC detection of any bacterium, fungus or parasite in a sample and for the
XX CC detection of at least one antimicrobial agent resistance gene or at
XX CC least one toxin gene. hexa nucleic acids are used for the specific and
XX CC ubiquitous detection and for identification of Streptococcus pneumoniae.
XX CC (1) can be used to design a therapeutic agent which is effective against
XX CC microorganisms. Microbial species or genus or family or phylum or group
XX CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
XX CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
XX CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,
XX CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
XX CC provides faster results than substrate specificity tests as results can
XX CC be determined in an hour and improved accuracy is also achieved.
XX CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
XX CC which are given in the exemplification of the present invention.
XX
XX
SQ Sequence 2157 BP; 696 A; 494 C; 405 G; 562 T; 0 other;

Query Match 15.9%; Score 318; DB 22; Length 2157;
Best Local Similarity 99.0%; Pred. No. 1.2e-153;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TAAATCTACGACAAATAAATCACTGCTGACTTGGGTTCTGAACGCCGCTCAA 60
DB 159 taaatctacgacataaaatcaatcattgctgacttggtctgaagcgcgctcaa 218
QY 61 TGCCCAAGCTAATGATATTCACAGATTTGGTTAAGCAATCGTTCTATCGAAGACA 120

```
DB 219 tgcgaagctaatgatattcccaagatttggtlaagcaatcgtttctatcgaagacca 278
QY 121 TCGCTTCTTCGACNACAGGGGATTTATACATCCGTATCCCTGGAGCTTCTGGGCAA 180
DB 279 tgcgtcttcgcacaaaggaggatltgattccatccgtatccctggagcttcttgagcaa 338
QY 181 TGTGAAGCAATTCCTCTCAAGGTGATCATCTCACCACCAAGTATTAAAGTTGAC 240
DB 339 tctgcaagtaattccctccaagtgatcaactccaccacaagtttaagttagac 398
QY 241 TTACTTTTCACTTCTCGACACGACTATTCTGTAAGGCTCAGGAAGCTTGTT 300
DB 399 ttactttcaactcgactccgaccagactattctcgtlaagctcagaagcttggtt 458
QY 301 AGCGATTGCTAGTAAGCAAAAGCAACGACAGAAATCTGACTATATATAATA 360
DB 459 agcgattcagtttagaacaagaacaaagaaatcttgaacctataataataa 518
QY 361 GGTCTACATGCTTAATGGGAACTATGGAATGACAGACAGCTCAAAACTATGTA 420
DB 519 ggtctacatgcttaattggaactatggaatgacagacagctcaaaactatagttaa 578
QY 421 AGACTCAATTAATTAAGTTTACCTCAGTTAGCTTGTGCTGGAATGCTCAGGAC 480
DB 579 agacctcaataatttaagttactcctagcttagcttgcgtggaatgctccagcacc 638
QY 481 AAACCAATATGACCCCTATTTCATTCAGACAGACCCGCAAGCCGCAAACTTGCTT 540
DB 639 aaaccaaatgacccctattccatccagaagcagcccaagccgcgaacttggctt 698
QY 541 ACTGTAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGGAAGAGCTCAATAC 600
DB 699 atctgaaatgaaataatcaagttactatctctctgtaacagttatgaaagcgtcaatc 758
QY 601 ACCAATTACTGATGAGATCAAAAGTCTCAAAATCAGCAAGTAATTACCTGTTACATGA 660
DB 759 accaatlactgattgactacaagaagtcacaatcagaagaagtaattacccttactgta 818
QY 661 TAATTAACCTCAGGA 675
DB 819 taattaccacaaga 833

RESULT 24
AAH01178
ID AAH01178 standard; DNA; 2160 BP.
XX
XX AAH01178;
XX
XX 24-JUL-2001 (first entry)
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1169.
XX
XX Species specific; genus specific; family specific; probe; detection;
XX Identification; algal; archaeal; bacterial; fungal; parasitical;
XX microorganism; diagnosis; translation elongation factor Tu; toxin;
XX translation elongation factor G; RecA recombinase; resistance;
XX catalytic subunit of proton-translocating ATPase; antimicrobial;
XX vaccine; primer; ds.
XX
XX Streptococcus pneumoniae.
XX
XX WO200123604-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000MO-CA01150.
XX
XX 28-SEP-1999; 99CA-2283458.
XX
XX 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
```

```
XX
PI Bergeron MC, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
XX WPI: 2001-245006/25.
XX
XX
XX Nucleic acid sequences are used to generate universal probes and
XX primers which can be used to identify and detect the presence of algal,
XX archaeal, bacterial, fungal and parasitical species in a test sample -
XX
XX Disclosure: Page 1044-1045; 1580pp; English.
XX
XX
XX The present invention describes a method for generating a repository of
XX nucleic acids of tuf, fus, atpD and/or recA genes from which probes
XX and/or primers are derived. The method comprises amplifying the nucleic
XX acids of determined algal, archaeal, bacterial, fungal and parasitical
XX species with a combination of defined primer pairs. The method can be
XX used for producing probes and/or primers for detecting one or more
XX related microorganisms e.g. algae, archaea, bacteria, fungi and
XX parasites, for universal detection and for specific and ubiquitous
XX detection and identification of an algal, archaeal, bacterial, fungal
XX and parasitical species, genus, family and group. A nucleic acid (I)
XX obtained using the method of the invention can be used for the universal
XX detection of any bacterium, fungus or parasite in a sample and for the
XX detection of at least one antimicrobial agent resistance gene or at
XX least one toxin gene. hexA nucleic acids are used for the specific and
XX ubiquitous detection and for identification of Streptococcus pneumoniae.
XX (I) can be used to design a therapeutic agent which is effective against
XX microorganisms. Microbial species or genus or family or phylum or group
XX can be detected include Abiotrophia adiacens, Bordetella sp.,
XX Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
XX Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
XX Neisseria gonorrhoeae and Staphylococcus sp. Using DNA based tests
XX provides faster results than substrate specificity tests as results can
XX be determined in an hour and improved accuracy is also achieved.
XX AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
XX which are given in the exemplification of the present invention.
XX
XX Sequence 2160 BP; 698 A; 494 C; 403 G; 565 T; 0 other;
XX
XX
XX Query Match 15.9%; Score 318; DB 22; Length 2160;
XX Best Local Similarity 99.0%; Pred. No. 1.2e-153;
XX Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX
XX 1 TAAATCTACGACATTAATAATCACTGATTTGCTGCTTGGTTGCAACGCCGCTCA 60
XX
XX 159 taaatctacgaacaataaaatcaactatgtgacttgggttctgaaacgcgcgtcaa 218
XX
XX 61 TGGCCAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTATCGAAAGACCA 120
XX
XX 219 tgcgaagctaatgatattcccaagatttggtlaagcaatcgtttctatcgaagacca 278
XX
XX 121 TCGCTTCTTCGACNACAGGGGATTTATACATCCGTATCCCTGGAGCTTCTGGGCAA 180
XX
XX 279 tgcgtcttcgcacaaaggaggatltgattccatccgtatccctggagcttcttgagcaa 338
XX
XX 181 TGTGAAGCAATTCCTCTCAAGGTGATCATCTCACCACCAAGTATTAAAGTTGAC 240
XX
XX 339 tctgcaagtaattccctccaagtgatcaactccaccacaagtttaagttagac 398
XX
XX 241 TTACTTTTCACTTCTCGACACGACTATTCTGTAAGGCTCAGGAAGCTTGTT 300
XX
XX 399 ttactttcaactcgactccgaccagactattctcgtlaagctcagaagcttggtt 458
XX
XX 301 AGCGATTGCTAGTAAGCAAAAGCAACGACAGAAATCTGACTATATATAATA 360
XX
XX 459 agcgattcagtttagaacaagaacaaagaaatcttgaacctataataataa 518
XX
XX 361 GGTCTACATGCTTAATGGGAACTATGGAATGACAGACAGCTCAAAACTATGTA 420
XX
XX 519 ggtctacatgcttaattggaactatggaatgacagacagctcaaaactatagttaa 578
```

```
QY 421 AGACCTCAATATTTAGTTTACCTCAGTTAGCCTTGCTGGTGAATGCCCTCAGCACC 480
CC |
CC |
CC |
DB 579 agaccctcaataatlaattagcttagcttagcttgctgctggaatgcctcaggaacc 638
QY 481 AAACCATATGACCCCTATTCACATCCAGAGAGAGCCCAAGACCCGCGAACTTGCTT 540
CC |
CC |
DB 639 aaaccataatgacccctattcacatccagaagcagcccaagcgcgcgaacttgctt 698
QY 541 ATGTGAATGAAAAATCAAGGCTACATCTCTGCTGACACATATGAGAAAGCATATAC 600
CC |
CC |
DB 699 atctgaatataaataatcaaggttacatctctgtcgaacagatagagaagcagtcatac 758
QY 601 ACCAATTACTGATGAGTACAAAGTCTCAATATGACGAAATTAATCCCTGTTACATGA 660
CC |
CC |
DB 759 accaattactgactgactacaagttcacaatcagcaagtaattaccctgcttaccatg 818
QY 661 TAATTACCTCAGGA 675
DB 819 taattacctcaagga 833

RESULT 25
AAH02056
ID AAH02056 standard; DNA; 2160 BP.
XX
XX AAH02056;
AC
XX
XX 24-JUL-2001 (first entry)
DT
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2049.
XX
XX Species specific; genus specific; family specific; probe; detection;
XX identification; algal; archaeal; bacterial; fungal; parasitical;
XX microorganism; diagnosis; translation elongation factor Tu; toxin;
XX translation elongation factor G; RecA recombinase; resistance;
XX catalytic subunit of proton-translocating ATPase; antimicrobial;
XX vaccine; primer; ds.
XX
XX Streptococcus pneumoniae.
OS
XX
XX MO200123604-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 28-SEP-2000; 2000MO-CA01150.
PF
XX
XX 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
PA
XX
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
PI
XX
XX WPI; 2001-245006/25.
DR
XX
XX Nucleic acid sequences are used to generate universal probes and
XX primers which can be used to identify and detect the presence of algal,
XX archaeal, bacterial, fungal and parasitical species in a test sample -
XX
XX
XX Disclosure; Page 1464-1465; 1580pp; English.
PS
XX
XX
XX The present invention describes a method for generating a repertory of
XX nucleic acids of tuf, fus, atpd and/or recA genes from which probes
XX and/or primers are derived. The method comprises amplifying the nucleic
XX acids of determined algal, archaeal, bacterial, fungal and parasitical
XX species with a combination of defined primer pairs. The method can be
XX used for producing probes and/or primers for detecting one or more
XX related microorganisms e.g. algae, archaea, bacteria, fungi and
XX parasites, for universal detection and for specific and ubiquitous
XX detection and identification of an algal, archaeal, bacterial, fungal
XX and parasitical species, genus, family and group. A nucleic acid (I)
XX obtained using the method of the invention can be used for the universal
```

```
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexA nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
XX Sequence 2160 BP; 696 A; 495 C; 402 G; 567 T; 0 other;
```

```
Query Match 15.9%; Score 318; DB 22; Length 2160;
Best Local Similarity 99.0%; Pred. No. 1,2e-153;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 1 TAAATCTAGACATAAATAATCAACTCATTCGTCGATGGTGGTCAAGCCGCGCA 60
DB 159 taaatctagacaataaataaataaactcatctgacttggtctgaacgcgcgca 218
QY 61 TGCCCAAGCTAATGATATTCCTCCACAGATTGGTTAAGGCATCGTTCTATGAGACCA 120
DB 219 tgcccaagctaatgatatctccacagatttggttaaggaactgcttctatgagacca 278
QY 121 TCGCTTTTGACACACAGGGGATTTGATACCTCCGATCCCTGGAGCTTTCTTGCCAA 180
DB 279 tcgcttttgacacacaggggatttgatccatcgcgtatccttggagcttcttcgcgaa 338
QY 181 TCTGCAAGCAATTCCTCCAGGTGATCACTCAACCAACAGTTGATTAAGTTGAC 240
DB 339 tcgcaagaatctcctccaggtgatcactcaacccaagtattgataagttgac 398
QY 241 TTACTTTCAACTTCGACTTCGACCAAGTATTTCTGTAAGGCTCAGGAAGCTTGTT 300
DB 399 ttactttcaacttcgacttcgcacagactatttccgtaaggtctcagaagttggt 458
QY 301 AGCGATTTCAGTTAGACAAAAGCACCAAGCAAGAAATCTTGACTCTATATATAATA 360
DB 459 agcgatttcagttagacaaaagcaccaagcaagaaatcttgactctatataataa 518
QY 361 GGCTACATGCTAATGGAAGTATGATGACAGACAGCAGCCTCAAAACTATGTTAA 420
DB 519 ggcctacatgctaatggaaagctatgaaatgacagacagcctcaaaactactatgtaa 578
QY 421 AGACCTCAATATTTAGTTTACCTCAGTTAGCCTTGCTGGTGAATGCCCTCAGCACC 480
DB 579 agaccctcaataatlaattagcttagcttagcttgctgctggaatgcctcaggaacc 638
QY 481 AAACCATATGACCCCTATTCACATCCAGAGAGAGCCCAAGACCCGCGAACTTGCTT 540
DB 639 aaaccataatgacccctattcacatccagaagcagcccaagcgcgcgaacttgctt 698
QY 541 ATGTGAATGAAAAATCAAGGCTACATCTCTGCTGACACATATGAGAAAGCATATAC 600
DB 699 atctgaatataaataatcaaggttacatctctgtcgaacagatagagaagcagtcatac 758
QY 601 ACCAATTACTGATGAGTACAAAGTCTCAATATGACGAAATTAATCCCTGTTACATGA 660
DB 759 accaattactgactgactacaagttcacaatcagcaagtaattaccctgcttaccatg 818
QY 661 TAATTACCTCAGGA 675
DB 819 taattacctcaagga 833
```

```
RESULT 26
AA235946
```

ID	AA35946 standard; DNA; 1260 BP.
XX	
AC	AA35946;
XX	
DT	07-FEB-2000 (first entry)
XX	
DE	Streptococcus pneumoniae pbp1a TER isolate h) nucleotide sequence.
XX	
KW	Streptococcus pneumoniae; penicillin binding protein; pbp2b; pbp1a;
KW	transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KW	detection; identification; pneumococcal meningitis; ss.
XX	
OS	Streptococcus pneumoniae.
XX	
PN	ZA9807024-A.
XX	
PD	28-APR-1999.
XX	
PF	05-AUG-1998; 982A-0007024.
XX	
PR	01-AUG-1997; 972A-0006886.
XX	
PA	(SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA	(DUWI-) UNIV WITWATERSRAND.
PA	(MED-) MEDICAL RES COUNCIL.
XX	
PI	Klugman KP, Smith AM, Du Plessis M;
XX	
DR	WPI; 1999-601770/51.
XX	
PT	Polymerase chain reaction assays for detecting Streptococcus pneumoniae
PT	useful for the diagnosis of pneumococcal meningitis
XX	
PS	Claim 11; Fig 4; 63pp; English.

CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of *Streptococcus pneumoniae*
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC *S. pneumoniae*, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant *S. pneumoniae* strains in a sample.
CC The methods can be used for detecting *S. pneumoniae* strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be used to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC *S. pneumoniae* with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a *Streptococcus*
CC *pneumoniae* *pp1a* transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX

Query Match	13.3%	Score 266;	DB 20;	Length 1260;
Best Local Similarity	99.1%	Pred. No. 8.3e-127;		
Matches 466; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

QY	575	GAAGTGTGGAAAGCAGCTAATACACCAATTCTGTATGAGTACACAAATCTCAATCA	634
Db	1	gaacagctatcgaaagcgcgtcaaacaccaattacgcgacgacaaagctccaatca	60
QY	635	GCAGTAATTAATCCCTGCTTACATGGATAATTACCTCAAGGAGATCATCAATCAAGTTGAA	694
Db	61	gaaagtaattacctgcgtcttaacatgtaattacctcaaggaagcatcacaatgaattgaa	120
QY	695	GAAAGAAACAGGCTTAACTCTACAACTGGATGGATGTCTACACAAATGTAGACCA	754
Db	121	gaagaaacaggaataaacacggtccacaacgtaggacgacgctcacacaatgtagaccaa	180
QY	755	GAAGCTCAAAAAACATCTGTGGATATTTCACATCAACACGAAATACGTTGGCTATCCGAC	814

Db	181	gaagctcaaaaaacatctgtggtgaattattacaatacagaatacgtgtgctatccaag	240
Qy	815	GATGAATTTGCAAGTGGCTTGTACCACTGGTGATGTTTCTTAACGGTAAAGTCATTGCCAG	874
Db	241	gatgaattgcaagtcgcgtcttcacatgtgtatgttcttaacgylaaagcatctggccag	300
Qy	875	CTAGAGGACGCGCATCACTCAAGTAATGTTTCCCTTGGGAFTTAACAGAGATGTAAGACA	934
Db	301	ctagaagacagccatcagatcaagtaatggttcccttggaattaaccaagaagtggaaca	360
Qy	935	AACCGCGACTGGGGATCACTATGTAAGAACGATCAGACTATGCTCCTGCTTGGAGTAC	994
Db	361	aacgcgcagctgggagtcacatgaacagatcaacagatcatgctcgtcccttggaatc	420
Qy	995	GGTGTCTACGATTCACATGCTACTATGCTTTACAGATGAGCCCTTAACATA	1044
Db	421	gtgtgtcagatctcaactgctactatcgtttcaagatgagccctatacta	470

RESULT	27
AAH02059	
ID	AAH02059 standard; DNA; 1195 BP
AC	
XX	AAH02059;
XX	
DT	24-JUL-2001 (first entry)

Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2052.

KM Species specific; genus specific; family specific; probe; detection
 KM Identification; algal; archaeal; bacterial; fungal; parasitical;
 KM microorganism; diagnosis; translation elongation factor Tu; toxin;
 KM translation elongation factor G; RecA recombinase; resistance;
 KM catalytic subunit of proton-translocating ATPase; antimicrobial;
 KM vaccine; primer; ds.

Streptococcus pneumoniae.

PN WO200123604-A2

PD 05-APR-2001.

28-SEP-2000; 2000WO-CA01150.

28-SEP-1999: 99CA-22283458.

PR 19-MAY-2000; 2000CA-230/0110:
XX

PA (INF-E-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
XX

PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M,
PI Picard ET, Roy PH:

XX
DDT, 2001-24E006/2E

XX
XX

PT primers which can be used to identify and detect the presence of algal,

XX

The present invention describes a method for generating a repertory of nucleic acids of tul, fus, alpd and/or recA genes from which probes and/or primers are derived. The method comprises amplifying the nucleic acids of determined algal, archaeal, bacterial, fungal and parasitical species with a combination of defined primer pairs. The method can be used for producing probes and/or primers for detecting one or more related microorganisms e.g. algae, archaea, bacteria, fungi and CC parasites, for universal detection and for specific and ubiquitous detection and identification of an algal, archaeal, bacterial, fungal and parasitical species, genus, family and group. A nucleic acid (1) obtained using the method of the invention can be used for the universal detection of any bacterium, fungus or parasite in a sample and for the detection of at least one antimicrobial agent resistance gene or at

CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SQ Sequence 1195 BP; 385 A; 268 C; 223 G; 319 T; 0 other;

Query Match 11.4%; Score 228; DB 22; Length 1195;
Best Local Similarity 100.0%; Pred. No. 3.5e-107;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1767 GAATGGAGAAATTCGTAATTAATAATGATGCTGCTGACGTGGAACACACCTGCTCCAC 1826
|||||
Db 968 gaattggagaattcgtatttaataatgctgcttctagtggaactaccctgcacac 1027
- 1827 AACACCCCATCACTCACTGAAAGTTCAAGCTCATCATCAGATGTTCACTTCAAGTCTA 1886
|||||
Db 1028 acaacaccccatcaactgaaagtcaagctcattcattcagatgattcaacttcacagctta 1087
OY 1887 GCTAACACCTCCCAAGCACAATTAATAGTACGACTACCAATCCCAATTAATAGCGAAC 1946
|||||
Db 1088 gctaacacactcccaagacaataatagctactaccacataataataatagcgaac 1147
OY 1947 AATCAATATACACCCCTGATCAACAAATCAGATCCTCAACGACGAC 1994
|||||
Db 1148 aatcaaatataacccctgatacaacaataatcagaatcctcaaccgcacac 1195

RESULT 28
AAZ35944
ID AAZ35944 standard; DNA; 1260 BP.
XX
AC AAZ35944;
XX
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbp1a TER isolate f) nucleotide sequence.
XX
KW Streptococcus pneumoniae; penicillin binding protein; pbp2b; pbp1a;
KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KW detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.

XX
PN ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 972A-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UYWI-) UNIT WITWATERSRAND.
PA (MED-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX
DR WPI; 1999-601770/51.
XX

PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX
PS Claim 11; Fig 4; 63pp; English.

XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2b (pbp2b) gene
CC and the pbp1a gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1a transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 403 A; 291 C; 267 G; 299 T; 0 other;

Query Match 10.5%; Score 209; DB 20; Length 1260;
Best Local Similarity 99.4%; Pred. No. 2.2e-97;
Matches 309; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 734 GTCATACAAATGTATGACACCAAGACTCAAAACATCTGTGGATATTATACATACAGAC 793
|||||
Db 160 gtctacacaatgtatgacacaagaagctcaaaaacatctgtggatatttacaatacagac 219
OY 794 GAATACGTTGCTTATCCAGACGATGAATGCAAGTCCCTTACCATTTGATGTTCT 853
|||||
Db 220 gaatacgttgcttataccagaagatgaatgcaagctccttaccattgtgatgtctt 279
OY 854 AAGGTAAGTCATGCGCCAGCTAGGACACGCCATCAGTAAGTATGTTCTCTTGGGA 913
|||||
Db 280 aacgtaagtcaltgcccagctagagcagcaccatgaatgaatgcttccctcgga 339
OY 914 ATTACCAAGCAGTATGACCAAAACCGGACCTGGGATCACTATGAACCGATCAGAC 973
|||||
Db 340 attaccaagcagtggaacaacaacgcgactgggactcaactatgaacacgatacagac 399
OY 974 TATGCTCTGCTTGGAGTACGTTGCTTACGATTCACACTGCTACTATGCTTACGATGAG 1033
|||||
Db 400 tatgctctgcttggaatacgtgtctacgattcaactgctactatcgttcaacgacgag 459
OY 1034 CCCATATACATA 1044
|||||
Db 460 cccataacta 470

RESULT 29
AAH02063
ID AAH02063 standard; DNA; 1212 BP.
XX
AC AAH02063;
XX
DI 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2056.
XX
KW Streptococcus pneumoniae;
KW identification; genus specific; family specific; probe; detection;
KW microorganism; algal; archaeal; bacterial; fungal; parasitical;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.

XX
PN WO200123604-A2.
XX
PD 05-APR-2001.
XX

PF 28-SEP-2000; 2000WO-CA01150.
 XX
 PR 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 XX
 XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX WPI; 2001-245006/25.
 DR
 XX Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitological species in a test sample -
 XX
 XX Claim 27; Page 1470; 1580pp; English.
 PS
 XX The present invention describes a method for generating a repository of
 CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitological
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitological species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SO Sequence 1212 BP; 381 A; 292 C; 244 G; 295 T; 0 other;

Query Match 8.9%; Score 178; DB 22; Length 1212;
 Best Local Similarity 99.6%; Pred. No. 2.3e-81;
 Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1739 GATTGGAATATACCAAGGAGGCTCTACAGAAATGGAATTCGATTTAAATGCTGCT 1798
 DB 984 gatgtgaatataccagaggggctctacagaaatggaatctgtatttaaaatggtgtct 1043
 OY 1799 CATTACAGTGAACACCTGCTCCACAAACCCCACTCAAGTCAAGTTCACCTCA 1858
 DB 1044 cgtctcagctgagagctcactgctccacaacaccccccatcaactgaaatgctcaagctca 1103
 OY 1859 TCATCAGATAGTTCAACCTTCACAGTCTAGCTCAACCACTTCAGCAAGCAAAATATAGTACG 1918
 DB 1104 tcatcagatagttcaactcactcagctcagctcaacaccccaagcaataatagtagc 1163
 OY 1919 ACTACCAATCTCAACATATACGCAACCAATCAAAATCAACCCCTATC 1967
 DB 1164 actaccatctcacaataataacgcaacatcaataacacccctgctc 1212

RESULT 30
 AAH01025
 ID AAH01025 standard; DNA; 1222 BP.
 XX
 AC AAH01025;

XX 24-JUL-2001 (first entry)
 DT Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1016.
 XX
 DE Species specific; genus specific; family specific; probe; detection;
 XX identification; algal; archaeal; bacterial; fungal; parasitological;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds;
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200123604-A2.
 XX
 PD 05-APR-2001.
 XX
 XX 28-SEP-2000; 2000WO-CA01150.
 PF 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 XX
 XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX WPI; 2001-245006/25.
 DR
 XX Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitological species in a test sample -
 XX
 XX Claim 27; Page 969-970; 1580pp; English.
 PS
 XX The present invention describes a method for generating a repository of
 CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitological
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitological species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SO Sequence 1222 BP; 389 A; 270 C; 233 G; 330 T; 0 other;

Query Match 8.9%; Score 178; DB 22; Length 1222;
 Best Local Similarity 99.6%; Pred. No. 2.3e-81;
 Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1739 GATTGGAATATACCAAGGAGGCTCTACAGAAATGGAATTCGATTTAAATGCTGCT 1798
 DB 994 gatgtgaatataccagaggggctctacagaaatggaatctgtatttaaaatggtgtct 1053

QY 1799 GATTCTACGTGGAACTGACCTGCTCTCACACACACCCCATCACTGAAAGTTCAAGCTCA 1858
|||||
Db 1054 cgttctacgttgagctcaccctgcctcaacaaccccccatcactgaagtcaagctca 1113
QY 1859 TCATCAGATAGTTCACACTTTCACAGTACGTCACACCACTCCACACACAATATATAGTACG 1918
|||||
Db 1114 tcactcagatagttcaacttcacagctcagctcaccctccaagcaacaataatagtagc 1173
QY 1919 ACTACCAATCCTCAATCAATATACGCAACCAATCAATCAACCCCTGATC 1967
|||||
Db 1174 actaccatccttaacataatagcgaacatcaataacacccctgac 1222

RESULT 31
AAH01022
ID AAH01022 standard; DNA: 1220 BP.
XX
AC AAH01022;
XX
XX 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1013.
XX
KW Species specific; genus specific; family specific; probe; detection;
KW Identification; algal; archaeal; bacterial; fungal; parasitica;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO260123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
XX 19-MAY-2000; 2000CA-2307010.
XX

(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
PA Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
DR WPI: 2001-245006/25.
XX

Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitica species in a test sample -
XX
PS Claim 27; Page 967-968; 1580pp; English.

The present invention describes a method for generating a repertory of
CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitica
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitica species, genus, family and group. A nucleic acid (II)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexA nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Adenoviridae, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,

CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX

Sequence 1220 BP: 388 A; 269 C; 233 G; 330 T; 0 other;
SQ

Query Match 8.8%; Score 175; DB 22; Length 1220;
Best Local Similarity 99.6%; Pred. No. 8.1e-80;
Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1739 GATTGGAATATATACAGAGGCGCTCTACAGAAATGAGATTCGATTTAAATAGTGTCT 1798
|||||
Db 995 gatgggaataatacaccagaggggctctacagaatgaggaattcgtatlaaaatgggct 1054
QY 1799 GATTCTACGTGGAACTGACCTGCTCTCACACACACCCCATCACTGAAAGTTCAAGCTCA 1858
|||||
Db 1055 cgttctacgttgagctcaccctgcctcaacaaccccccatcactgaagtcaagctca 1114
QY 1859 TCATCAGATAGTTCACACTTTCACAGTACGTCACACCACTCCACACACAATATATAGTACG 1918
|||||
Db 1115 tcactcagatagttcaacttcacagctcagctcaccctccaagcaacaataatagtagc 1174
QY 1919 ACTACCAATCCTCAATCAATATACGCAACCAATCAATCAACCCCTG 1964
|||||
Db 1175 actaccatccttaacataatagcgaacatcaataacacccctg 1220

RESULT 32
AAH02069
ID AAH02069 standard; DNA: 1216 BP.
XX
AC AAH02069;
XX
XX 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2062.
XX
KW Species specific; genus specific; family specific; probe; detection;
KW Identification; algal; archaeal; bacterial; fungal; parasitica;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
XX 19-MAY-2000; 2000CA-2307010.
XX

(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
PA Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
DR WPI: 2001-245006/25.
XX

Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitica species in a test sample -
XX
PS Claim 27; Page 1474-1475; 1580pp; English.

The present invention describes a method for generating a repertory of

KW	catalyze	subunit of proton-translocating ATPase; antimicrobial;
XV	vaccine	primer; ds.
XX		
OS	Streptococcus pneumoniae.	
PN	MO200123604-AZ.	
XX		
XX	05-APR-2001.	
PD		
PF	28-SEP-2000; 2000MO-CA01150.	
XX		
PR	28-SEP-1999; 99CA-2283458-	
PR	19-MAY-2000; 2000CA-230710.	
PA	(INF-) INFECTIO DIAGNOSTIC (IDI) INC.	
XX		
PI	Bergeron MG, Boissinot M, Hulatsky A, Menard C, Ouellette M;	
PI	Picard EJ, Roy PH;	
XX		
DR	WPI: 2001-245006/25.	
XX		
PT	Nucleic acid sequences are used to generate universal probes and	
PT	primers which can be used to identify and detect the presence of algal,	
PT	archaeal, bacterial, fungal and parasitical species in a test sample -	
PS	Claim 27; Page 961-962; 1580pp; English.	
XX		
CC	The present invention describes a method for generating a repertory of	
CC	nucleic acids of tuf, fus, atpd and/or recA genes from which probes	
CC	and/or primers are derived. The method comprises amplifying the nucleic	
CC	acids of determined algal, archaeal, bacterial, fungal and parasitical	
CC	species with a combination of defined primer pairs. The method can be	
CC	used for producing probes and/or primers for detecting one or more	
CC	related microorganisms e.g. algae, archaea, bacteria, fungi and	
CC	parasites, for universal detection and for specific and ubiquitous	
CC	detection and identification of an algal, archaeal, bacterial, fungal	
CC	and parasitical species, genus, family and group. A nucleic acid (I)	
CC	obtained using the method of the invention can be used for the universal	
CC	detection of any bacterium, fungus or parasite in a sample and for the	
CC	detection of at least one antimicrobial agent resistance gene or at	
CC	least one toxin gene. hexA nucleic acids are used for the specific and	
CC	ubiquitous detection and for identification of Streptococcus pneumoniae.	
CC	(I) can be used to design a therapeutic agent which is effective against	
CC	microorganisms. Microbial species or genus or family or phylum or group	
CC	which can be detected include Abiotrophia adiacens, Bordetella sp.,	
CC	Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,	
CC	Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,	
CC	Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests	
CC	provides faster results than substrate specificity tests as results can	
CC	be determined in an hour and improved accuracy is also achieved.	
CC	AH00010 to AH002304 represent nucleotide sequences and primers/probes	
CC	which are given in the exemplification of the present invention.	
XX		
SQ	Sequence 1212 BP; 391 A; 265 C; 225 G; 321 T; 0 other:	
	Query Match	6.4%; Score 127; DB 22; Length 1212;
	Best Local Similarity	99.1%; Pred. No. 4, 9e-55;
	Matches 227; Conservative	0; Mismatches 2; Indels 0; Gaps 0
OY	1739 GATTGGAAATATACGAGAGGGGCTCTACACGAATAATGGAGATTCGTATTAAAAATGGTGGCT	1798
Dd	gatttgaataatccagaaagggtctctacaagaatggagtctgtatttaaatggttgtc	1043
OY	1799 GATTTCAGCTGGAACGCACACTCGCTCCACACAACCCTCATCACAAGTTTCACACTCA	1858
Dd	gatttcacgttgagcgcacctcgctccacaacaaccctcacactgaaaagttcaagctca	1103
OY	1859 TCATCAGAATAGTTCAACTCACAGTCTACTACCAACCATCTCAAGCAACAATATAATAGTAGC	1918
Dd	tcatcagaatagttcaacttcacagctctaactcaaccctccaagcacaataatatagtagc	1163
OY	1919 ACTACCAATCTCAACATATATACGACAACAATCAATCAACCCCTTGATC	1967

Db 1164 attaccatcctaacaataatcgcacaataatcacaaccctgac 1212

|||||

RESULT 35

AAH01017

ID AAH01017 standard; DNA; 1223 BP.

XX

AC AAH01017;

XX

DT 24-JUL-2001 (first entry)

XX

DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1008.

XX

Species specific; genus specific; family specific; probe; detection;

KM identification; algal; archaeal; bacterial; fungal; parasitic;

KM microorganism; diagnosis; translation elongation factor Tu; toxin;

KW translation elongation factor G; RecA recombinase; resistance;

KW catalytic subunit of proton-translocating ATPase; antimicrobial;

KW vaccine; primer; ds.

XX

Streptococcus pneumoniae.

OS

PN WO200123604-A2.

XX

PD 05-APR-2001.

XX

PF 28-SEP-2000; 2000MO-CA01150.

XX

PR 28-SEP-1999; 99CA-2283458.

PR 19-MAY-2000; 2000CA-2307010.

XX

(INF-) INFECTION DIAGNOSTIC (IDI) INC.

PA Bergeron MG, Boltsinot M, Huletsky A, Menard C, Ouellette M;

PI Picard FJ, Roy PH;

DR WPI: 2001-245006/25.

XX

Nucleic acid sequences are used to generate universal probes and

PT primers which can be used to identify and detect the presence of algal,

PT archaeal, bacterial, fungal and parasitic species in a test sample -

PS Claim 27: Page 964; 1580bp; English.

XX

The present invention describes a method for generating a repertoire of

CC nucleic acids of tuf, fus, apd and/or recA genes from which probes

CC and/or primers are derived. The method comprises amplifying the nucleic

CC acids of determined algal, archaeal, bacterial, fungal and parasitic

CC species with a combination of defined primer pairs. The method can be

CC used for producing probes and/or primers for detecting one or more

CC related microorganisms e.g. algae, archaea, bacteria, fungi and

CC parasites, for universal detection and for specific and ubiquitous

CC detection and identification of an algal, archaeal, bacterial, fungal

CC and parasitic species, genus, family and group. A nucleic acid (1)

CC obtained using the method of the invention can be used for the universal

CC detection of any bacterium, fungus or parasite in a sample and for the

CC detection of at least one antimicrobial agent resistance gene or at

CC least one toxin gene. hexa nucleic acids are used for the specific and

CC ubiquitous detection and for identification of Streptococcus pneumoniae.

CC (1) can be used to design a therapeutic agent which is effective against

CC microorganisms. Microbial species or genus or family or phylum or group

CC which can be detected include Abiotrophia adiacens, Bordetella sp.,

CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,

CC Mycobacteriaceae family, Pseudomonas group, Streptococcus sp.,

CC Neisseria gonorrhoeae and Staphylococcus sp. Using DNA based tests

CC provides faster results than substrate specificity tests as results can

CC be determined in an hour and improved accuracy is also achieved.

CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes

CC which are given in the exemplification of the present invention.

XX

Sequence 1223 BP; 394 A; 271 C; 235 G; 323 T; 0 other;

PT archaeal, bacterial, fungal and parasitical species in a test sample -
 XX
 PS Claim 27: Page 963; 1580bp; English.
 XX
 CC The present invention describes a method for generating a repertory of
 CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitical
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitical species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (1) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SQ Sequence 1218 BP; 392 A; 271 C; 233 G; 322 T; 0 other;

Query Match 6.3%; Score 126; DB 22; Length 1218;
 Best Local Similarity 100.0%; Pred. No. 1.6e-54;
 Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1842 CTGAAGTTCAGCTCATCATGATGTTCACTTACAGCTTACGCTCAACCACTCCAA 1901
 |||||||
 DB 1093 ctgaagttcaagctcatcatcagatagttcaacttcacagtcacacccactccaa 1152

QY 1902 GCACAAATATAGTACGACTCCAACTTAATATACCAATATCAATATCAACCC 1961
 |||||||
 DB 1153 gcacaataatagtcagactaacctccataataacgacaatacaataacc 1212

QY 1962 CTGATC 1967
 |||||||
 DB 1213 ctgctc 1218

RESULT 38
 AAH01015
 ID AAH01015 standard; DNA: 1213 BP.
 XX
 AC AAH01015;
 XX
 DT 24-JUL-2001 (first entry)
 XX
 DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1006.
 XX
 XX Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitical;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200123604-A2.
 XX
 PD 05-APR-2001.
 XX

PF 28-SEP-2000; 2000WO-CA01150.
 XX
 PR 28-SEP-1999; 99CA-2283458.
 PR 19-MAY-2000; 2000CA-2307010.
 XX
 XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
 PA Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;
 XX
 DR WPI: 2001-24506/25.
 XX
 XX Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitical species in a test sample -
 XX
 PS Claim 27: Page 962-963; 1580bp; English.

The present invention describes a method for generating a repertory of
 CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitical
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitical species, genus, family and group. A nucleic acid (II)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (1) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.
 XX
 SQ Sequence 1213 BP; 393 A; 266 C; 235 G; 319 T; 0 other;

Query Match 6.1%; Score 122; DB 22; Length 1213;
 Best Local Similarity 99.1%; Pred. No. 1.9e-52;
 Matches 222; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1739 GATTGGAATATATACAGAGGGCTCTACAGAAATGAGAAATCGTATTTAAATGGTCT 1798
 |||||||
 DB 990 gattggaatatacagagggctctacagaaatgagaaatcgtaatttaaaatgggtct 1049

QY 1799 CGTCTAGTGAACCTCACTGCTCCACACACACCCCATCATCACTGAAAGTTCAAGTCA 1858
 |||||||
 DB 1050 cgtctacgttgaggtcaccgctccacacacaccccatcaactgaagtcaagctca 1109

QY 1859 TCATCAGATAGTCAACTTCACTGCTTCAAGTCAACCACTCCACAGACAAATATAGTACG 1918
 |||||||
 DB 1110 tcatcagtagttcaacttcacagtcctcaaccactccaagacacaataatagtcag 1169

QY 1919 ACTACCAATCTCAACATATATACGACACCAATCAATATCAACCC 1962
 |||||||
 DB 1170 actaccaatctcaacatataatatacgaacaataatacaacc 1213

RESULT 39
 AAH01018
 ID AAH01018 standard; DNA: 1214 BP.
 XX
 AC AAH01018;

XX 24-JUL-2001 (first entry)
XX
XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1009.
DE
XX Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitital;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
XX Streptococcus pneumoniae.
OS
XX WO200123604-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-CA01150.
XX
XX 28-SEP-1999; 99CA-2283458.
XX 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
XX WPI: 2001-245006/25.
XX
XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitital species in a test sample -
XX
XX Claim 27; Page 964-965; 1580pp; English.
XX
XX The present invention describes a method for generating a repertory of
XX nucleic acids of tuf, fus, atp and/or recA genes from which probes
XX and/or primers are derived. The method comprises amplifying the nucleic
XX acids of determined algal, archaeal, bacterial, fungal and parasitital
XX species with a combination of defined primer pairs. The method can be
XX used for producing probes and/or primers for detecting one or more
XX related microorganisms e.g. algae, archaea, bacteria, fungi and
XX parasites, for universal detection and for specific and ubiquitous
XX detection and identification of an algal, archaeal, bacterial, fungal
XX and parasitital species, genus, family and group. A nucleic acid (I)
XX obtained using the method of the invention can be used for the universal
XX detection of any bacterium, fungus or parasite in a sample and for the
XX detection of at least one antimicrobial agent resistance gene or at
XX least one toxin gene. hexA nucleic acids are used for the specific and
XX ubiquitous detection and for identification of Streptococcus pneumoniae.
XX (I) can be used to design a therapeutic agent which is effective against
XX microorganisms. Microbial species or genus or family or phylum or group
XX which can be detected include Abiotrophia adiacens, Bordetella sp.,
XX Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
XX Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
XX Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
XX provides faster results than substrate specificity tests as results can
XX be determined in an hour and improved accuracy is also achieved.
XX AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
XX CC that are given in the exemplification of the present invention.
XX
XX Sequence 1214 BP; 391 A; 270 C; 232 G; 321 T; 0 other;

Query Match 6.1%; Score 122; DB 22; Length 1214;
Best Local Similarity 100.0%; Pred. No. 1.9e-52;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1842 CTGAAGTTCACGCATCATCATGATGATGTCACATCTACAGCTTACTGCAACACCTCCAA 1901
DB 1093 ctgaaagttcaagctcatcatcatagcttagtcaacttcacagcttagctcaaccattccaa 1152

QY 1902 GCACAATATATAGTACGACTACCATCTCTACATATATAGCACAATCAATACACCC 1961
DB 1153 gcaacaataatagtagactaccatccatccataataataatagcaacaataataacc 1212
QY 1962 CT 1963
DB 1213 ct 1214
RESULT 40
AAH01020
ID AAH01020 standard; DNA; 1207 BP.
XX
XX AAH01020;
AC
XX 24-JUL-2001 (first entry)
XX
XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1011.
DE
XX Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitital;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
XX Streptococcus pneumoniae.
OS
XX WO200123604-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-CA01150.
XX
XX 28-SEP-1999; 99CA-2283458.
XX 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
XX WPI: 2001-245006/25.
XX
XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitital species in a test sample -
XX
XX Claim 27; Page 966; 1580pp; English.
XX
XX The present invention describes a method for generating a repertory of
XX nucleic acids of tuf, fus, atp and/or recA genes from which probes
XX and/or primers are derived. The method comprises amplifying the nucleic
XX acids of determined algal, archaeal, bacterial, fungal and parasitital
XX species with a combination of defined primer pairs. The method can be
XX used for producing probes and/or primers for detecting one or more
XX related microorganisms e.g. algae, archaea, bacteria, fungi and
XX parasites, for universal detection and for specific and ubiquitous
XX detection and identification of an algal, archaeal, bacterial, fungal
XX and parasitital species, genus, family and group. A nucleic acid (I)
XX obtained using the method of the invention can be used for the universal
XX detection of any bacterium, fungus or parasite in a sample and for the
XX detection of at least one antimicrobial agent resistance gene or at
XX least one toxin gene. hexA nucleic acids are used for the specific and
XX ubiquitous detection and for identification of Streptococcus pneumoniae.
XX (I) can be used to design a therapeutic agent which is effective against
XX microorganisms. Microbial species or genus or family or phylum or group
XX which can be detected include Abiotrophia adiacens, Bordetella sp.,
XX Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
XX Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
XX Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
XX provides faster results than substrate specificity tests as results can
XX be determined in an hour and improved accuracy is also achieved.

CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SQ Sequence 1207 BP; 388 A; 268 C; 231 G; 320 T; 0 other;

Query Match 6.1%; Score 121; DB 22; Length 1207;
Best Local Similarity 100.0%; Pred. No. 6.2e-52;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1842 CTGAAAGTTCAGCTCATCATAGTTCACACTTCACTTCAAGCTCAACCACTCCAA 1901
|
DB 1087 ctgaagaagtcacagtcacacagatagttcaacttcacagtcacagtcacacccaa 1146
|
QY 1902 GCACAATTAATAGTACGACCTCAATCTTAATTAATAGCACAATCAATCAACCC 1961
|
DB 1147 gcacaataatagtagacacacacacacacacacacacacacacacacacacac 1206
|
QY 1962 C 1962
|
DB 1207 c 1207

RESULT 41

AAH01013
ID AAH01013 standard; DNA; 1212 BP.

XX
AC AAH01013;

XX
DT 24-JUL-2001 (first entry)

XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1004.

XX
KM Species specific; genus specific; family specific; probe: detection;
KM identification: algal; archaeal; bacterial; fungal; parasitica;
KM microorganism; diagnosis; translation elongation factor Tu; toxin;
KM translation elongation factor G; RecA recombinase; resistance;
KM catalytic subunit of proton-translocating ATPase; antimicrobial;
KM vaccine; primer; ds.

XX
OS Streptococcus pneumoniae.

XX
PN WO200123604-A2.

XX
PD 05-APR-2001.

XX
PF 28-SEP-2000; 2000WO-CA01150.

XX
PR 28-SEP-1999; 99CA-2283458.

XX
PR 19-MAY-2000; 2000CA-2307010.

XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.

XX
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;

XX
PI Picard FJ, Roy PH;

XX
DR WPI; 2001-245006/25.

XX
PT Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitica species in a test sample -
XX
PS Claim 27; Page 961; 1580pp; English.

XX
CC The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitica
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific, bacterial, fungal
CC detection and identification of an algal, archaeal, bacterial, fungal

CC and parasitica species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SQ Sequence 1212 BP; 391 A; 269 C; 232 G; 320 T; 0 other;

Query Match 6.0%; Score 120; DB 22; Length 1212;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1842 CTGAAAGTTCAGCTCATCATAGTTCACACTTCACTTCAAGCTCAACCACTCCAA 1901
|
DB 1093 ctgaagaagtcacagtcacacagatagttcaacttcacagtcacagtcacacccaa 1152
|

QY 1902 GCACAATTAATAGTACGACCTCAATCTTAATTAATAGCACAATCAATCAACCC 1961
|
DB 1153 gcacaataatagtagacacacacacacacacacacacacacacacacacacac 1212
|

RESULT 42

AAH02064
ID AAH02064 standard; DNA; 1242 BP.

XX
AC AAH02064;

XX
DT 24-JUL-2001 (first entry)

XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2057.

XX
KM Species specific; genus specific; family specific; probe: detection;
KM identification: algal; archaeal; bacterial; fungal; parasitica;
KM microorganism; diagnosis; translation elongation factor Tu; toxin;
KM translation elongation factor G; RecA recombinase; resistance;
KM catalytic subunit of proton-translocating ATPase; antimicrobial;
KM vaccine; primer; ds.

XX
OS Streptococcus pneumoniae.

XX
PN WO200123604-A2.

XX
PD 05-APR-2001.

XX
PF 28-SEP-2000; 2000WO-CA01150.

XX
PR 28-SEP-1999; 99CA-2283458.

XX
PR 19-MAY-2000; 2000CA-2307010.

XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.

XX
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;

XX
PI Picard FJ, Roy PH;

XX
DR WPI; 2001-245006/25.

XX
PT Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitica species in a test sample -
XX
PS Claim 27; Page 1471; 1580pp; English.

XX Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
KM transpeptidase encoding region; TBR: antibiotic resistance; diagnosis;
KW detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 98ZA-0007024.
XX
PR 01-AUG-1997; 97ZA-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
XX (UWMI-) UNIV WITWATERSRAND.
XX (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
DR WPI: 1999-601770/51.
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis -
XX
PS Claim 11; Fig 4; 63pp; English.
XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbp1A gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbp1A transpeptidase encoding region (TBR) isolate nucleotide
CC sequence from the present invention.
XX
SQ Sequence 1260 BP; 394 A; 264 C; 258 G; 344 T; 0 other;

Query Match 5.1%; Score 101; DB 20; Length 1260;
Best Local Similarity 100.0%; Pred. No. 1.3e-41;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 575 GAACAGTATGAGAAAGACGATACACCAATTACTGATGAGACTACAAAGTCTCAATCA 634
DB 1 gaacagatagagaagcagatcaatcaccaattactgactgagctacaagtcctcaatca 60'
QY 635 GGAAGTAATTACCGCTTACATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 675
DB 61 gcaagtaattaccctgcttaccatgataatcctcaaga 101

RESULT 47
AAH02057
ID AAH02057 standard; DNA; 930 BP.
XX
XX AAH02057;
AC
XX
DT 24-JUL-2001 (first entry)
XX
XX Streptococcus pneumoniae nucleotide sequence SEQ ID NO: 2050.
DE
XX
XX Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitical;

KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KM translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
XX PR 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
PI
XX
DR WPI: 2001-245006/25.
XX
XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitical species in a test sample -
XX
PS Disclosure; Page 1466; 1580pp; English.
XX
CC The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitical
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitical species, genus, family and group. A nucleic acid (I)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexA nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SQ Sequence 930 BP; 282 A; 195 C; 192 G; 261 T; 0 other;

Query Match 4.9%; Score 97; DB 22; Length 930;
Best Local Similarity 100.0%; Pred. No. 1.5e-39;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1377 AAAAAGTTCCTATATGTCGACACGTCGCTTGAAGAAAGCAAGCCTATATGATCA 1436
DB 608 aaaaagatctcttaatgctgagactcgtgcatgagaagaacagacctatagatga 667
QY 1437 CCGACATGATGAAAACAGCTTGTACTTATGAACTGG 1473
DB 668 ccgacatgatgaaaacagctctgtactatggaactgg 704

RESULT 48
AAZ35942

ID AA235942 standard; DNA: 1260 BP.
XX
AC AA235942;
XX
DT 07-FEB-2000 (first entry)
XX
DE Streptococcus pneumoniae pbpia TER isolate d) nucleotide sequence.
XX
DE Streptococcus pneumoniae: penicillin binding protein: pbp2B; pbpia:
KM transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
KW detection; identification; pneumococcal meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
PN ZA9807024-A.
XX
PD 28-APR-1999.
XX
PF 05-AUG-1998; 982A-0007024.
XX
PR 01-AUG-1997; 97ZA-0006886.
XX
PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.
PA (UYWI-) UNIV WITWATERSRAND.
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Klugman KP, Smith AM, Du Plessis M;
XX WPI; 1999-601770/51.
DR
XX
PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
PT useful for the diagnosis of pneumococcal meningitis
XX
XX
PS Claim 11; Fig 4; 63pp; English.
XX
CC A polymerase chain reaction (PCR) assays have been developed for
CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
CC using primers based on the penicillin binding protein 2B (pbp2B) gene
CC and the pbpia gene. The products and methods can be used for detecting
CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
CC used for simultaneously diagnosing pneumococcal meningitis and
CC identifying any antibiotic-resistant S. pneumoniae strains in a sample.
CC The methods can be used for detecting S. pneumoniae strains resistant
CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
CC The assays can be adapted to detect other pathogens causing meningitis.
CC The assays can be used to detect an antibiotic resistant strain of
CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
CC a 224 bp product. The present sequence represents a Streptococcus
CC pneumoniae pbpia transpeptidase encoding region (TER) isolate nucleotide
CC sequence from the present invention.
XX
SO Sequence 1260 BP; 404 A; 266 C; 254 G; 336 T; 0 other;

Query Match 4.9%; Score 97; DB 20; Length 1260;
Best Local Similarity 100.0%; Pred. No. 1.5e-39;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1377 AAAAAGATTCTATATGTCGAACTGTCGATGAAGAAGACAGCCTATATGATGA 1436
DB 803 aaaaagattctctaatatgctgaactcgtgcataagaagaacagcactatagtga 862
OY 1437 CCGACATGATGAAGAAAGCTGTGACTATANGAAGACTGG 1473
DB 863 ccgacatgatgaagaagctgtgactatgaactgg 899

RESULT 49
AAH02071
ID AAH02071 standard; DNA: 782 BP.
XX
AC AAH02071;

XX
DE 24-JUL-2001 (first entry)
XX
DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:2064.
XX
XX
KW Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitical;
KW microorganism; diagnosis; translation elongation factor Tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Streptococcus pneumoniae.
XX
PN W0200123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
XX WPI; 2001-245006/25.
DR
XX
PT Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitical species in a test sample -
XX
XX
PS Claim 27; Page 1476; 1580pp; English.
XX
CC The present invention describes a method for generating a repository of
CC nucleic acids of ruf, fus, atpd and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic
CC acids of determined algal, archaeal, bacterial, fungal and parasitical
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitical species, genus, family and group. A nucleic acid (II)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexa nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (I) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Adenoviridae, Bacteroides, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SO Sequence 782 BP; 247 A; 181 C; 156 G; 198 T; 0 other;

Query Match 4.5%; Score 90; DB 22; Length 782;
Best Local Similarity 100.0%; Pred. No. 6.2e-36;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 965 ATCAGACATATGCTCTGCTGCTTGGAGTACGGGTCTACAGTTAACTGCTACTATGTT 1024
DB 163 atcacagcatatgctctgctgcttggagtagtgctcagatcactgctactatgctt 222

QY 1025 CACGATGAGCCCTATTAACCTGGGACA 1054
 Db 223 cacgatgagccctataactaccctggaca 252

RESULT 50

AAH01655
 ID AAH01655 standard; DNA; 813 BP.

AC AAH01655;

DT 24-JUL-2001 (first entry)

DE Streptococcus pneumoniae nucleotide sequence SEQ ID NO:1648.

XX Species specific; genus specific; family specific; probe; detection;
 KW identification; algal; archaeal; bacterial; fungal; parasitical;
 KW microorganism; diagnosis; translation elongation factor Tu; toxin;
 KW translation elongation factor G; RecA recombinase; resistance;
 KW catalytic subunit of proton-translocating ATPase; antimicrobial;
 KW vaccine; primer; ds.

OS Streptococcus pneumoniae.

XX WO200123604-A2.

XX 05-APR-2001.

PD 28-SEP-2000; 2000WO-CA01150.

XX 28-SEP-1999; 99CA-2283458.

PR 19-MAY-2000; 2000CA-2307010.

XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.

XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
 PI Picard FJ, Roy PH;

XX WPI: 2001-245006/25.

PT Nucleic acid sequences are used to generate universal probes and
 PT primers which can be used to identify and detect the presence of algal,
 PT archaeal, bacterial, fungal and parasitical species in a test sample -

PS Claim 27: Page 1255; 1580pp; English.

CC The present invention describes a method for generating a repertory of
 CC nucleic acids of tuf, fus, atpd and/or recA genes from which probes
 CC and/or primers are derived. The method comprises amplifying the nucleic
 CC acids of determined algal, archaeal, bacterial, fungal and parasitical
 CC species with a combination of defined primer pairs. The method can be
 CC used for producing probes and/or primers for detecting one or more
 CC related microorganisms e.g. algae, archaea, bacteria, fungi and
 CC parasites, for universal detection and for specific and ubiquitous
 CC detection and identification of an algal, archaeal, bacterial, fungal
 CC and parasitical species, genus, family and group. A nucleic acid (I)
 CC obtained using the method of the invention can be used for the universal
 CC detection of any bacterium, fungus or parasite in a sample and for the
 CC detection of at least one antimicrobial agent resistance gene or at
 CC least one toxin gene. hexa nucleic acids are used for the specific and
 CC ubiquitous detection and for identification of Streptococcus pneumoniae.
 CC (I) can be used to design a therapeutic agent which is effective against
 CC microorganisms. Microbial species or genus or family or phylum or group
 CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
 CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
 CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
 CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
 CC provides faster results than substrate specificity tests as results can
 CC be determined in an hour and improved accuracy is also achieved.
 CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
 CC which are given in the exemplification of the present invention.

XX Sequence 813 BP; 253 A; 186 C; 164 G; 210 T; 0 other;

Query Match 4.5%; Score 90; DB 22; Length 813;
 Best Local Similarity 100.0%; Pred. No. 6,2e-36;
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 965 ATCAGACATATGCTCCTGCTTGAGTACGCTGTCTACGATTCACGCTACTATTCGTT 1024
 Db 169 atcacagactatgctctccctgcttgagtaagtgctacgattcaactactactatcggt 228
 QY 1025 CACGATGAGCCCTATTAACCTGGGACA 1054
 Db 229 cacgatgagccctataactaccctggaca 258

Search completed: June 13, 2002, 12:42:27
 Job time: 6595 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 08:39:33 : Search time 2129.67 Seconds

(without alignments)
12668.823 Million cell updates/sec

Title: US-08-961-083-1

Perfect score: 1999

Sequence: 1 TAAATCTAGACATATAAA.....ATCCTACACGACAAACA 1999

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 15

Total number of hits satisfying chosen parameters: 25133

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	1.2	928	10 BG334672	BG334672 602460969
2	22	1.1	553	12 AO701777	AO701777 HS_2134_A
3	22	1.1	909	12 CNS0363A	AL229519 Tetracodon
4	21	1.1	389	9 AL632585	AL632585 AL632585
5	21	1.1	422	9 AI397462	AI397462 FB1808.x
6	21	1.1	704	9 AL629478	AL629478 AL629478
7	21	1.1	1025	12 CNS06N80	AL406462 T3 end of
8	21	1.1	1078	12 CNS06NM7	AL406462 T7 end of
9	20	1.0	237	10 BF174868	BF174868 MYE3951 M
10	20	1.0	287	9 AA317576	AA317576 EST19524
11	20	1.0	295	9 BB551207	BB551207 BB551207
12	20	1.0	360	10 BF016990	BF016990 ux19d08.y
13	20	1.0	367	10 BF016981	BF016981 ux19c08.y
14	20	1.0	402	10 BM173589	BM173589 900943 Av
15	20	1.0	405	10 BF016982	BF016982 ux19c09.y
16	20	1.0	406	12 AQ445724	AQ445724 nbxb0054J
17	20	1.0	440	10 BG017796	BG017796 dab10f11.

18	20	1.0	442	9	BB824552	BB824552
19	20	1.0	443	10	BE566388	BE566388 601339984
20	20	1.0	446	10	BG942702	BG942702 ax28e04.x
21	20	1.0	450	9	BB750480	BB750480 BB750480
22	20	1.0	493	9	AA280963	AA280963 zs97d12.x
23	20	1.0	526	9	AW762116	AW762116 ur53h07.y
24	20	1.0	526	12	AO710693	AO710693 HS_5344_A
25	20	1.0	535	12	A2824983	A2824983 2M0099J2A
26	20	1.0	563	10	BG723460	BG723460 602694145
27	20	1.0	587	10	BG503218	BG503218 602651318
28	20	1.0	588	10	BG539230	BG539230 602657826
29	20	1.0	642	9	AA606845	AA606845 vmy94h07.r
30	20	1.0	650	10	BG492386	BG492386 602536313
31	20	1.0	661	9	BA498524	BA498524 BA498524
32	20	1.0	669	10	BF212414	BF212414 601813376
33	20	1.0	677	10	BG989708	BG989708 HOA56-1-G
34	20	1.0	705	9	AA152866	AA152866 mq56c03.r
35	20	1.0	710	10	BF658632	BF658632 maa98h01.
36	20	1.0	712	12	AG099717	AG099717 Pan trogl
37	20	1.0	731	10	BG506942	BG506942 601861569
38	20	1.0	737	12	A2354492	A2354492 1M0093T15
39	20	1.0	787	10	BF233667	BF233667 601906930
40	20	1.0	788	10	BE889059	BE889059 601513993
41	20	1.0	803	10	BI690754	BI690754 603314371
42	20	1.0	841	10	BF667800	BF667800 602122620
43	20	1.0	861	10	BE383813	BE383813 601298315
44	20	1.0	873	12	BH466047	BH466047 BCGV80TR
45	20	1.0	888	10	BF670847	BF670847 602149929
46	20	1.0	896	10	BF696798	BF696798 602125331
47	20	1.0	904	10	BF247101	BF247101 601854544
48	20	1.0	917	10	BF697265	BF697265 602129793
49	20	1.0	938	10	BG282996	BG282996 602406185
50	20	1.0	969	12	AO688268	AO688268 nbxb0077M
51	20	1.0	1021	10	BE964433	BE964433 601658140
52	20	1.0	1101	10	BM480145	BM480145 AGENCOURT
53	20	1.0	1151	10	BF797518	BF797518 602257187
54	20	1.0	1537	10	BM457339	BM457339 AGENCOURT
55	20	1.0	3519	11	AK010591	AK010591 Mus muscu
56	19	1.0	120	12	TA129D01P	TA129D01P
57	19	1.0	225	9	AV377407	AV377407
58	19	1.0	312	9	AV299242	AV299242
59	19	1.0	319	10	BG801896	BG801896 0122-44 M
60	19	1.0	325	12	AQ567904	AQ567904 HS_2117_A
61	19	1.0	367	9	AM861079	AM861079 PM4-CT029
62	19	1.0	375	10	BM264807	BM264807 fw62g01.x
63	19	1.0	401	10	Z29739	Z29739 ARTS2183 Or
64	19	1.0	407	10	BE925196	BE925196 CM2-AN007
65	19	1.0	408	9	BE039091	BE039091 AB09E06 A
66	19	1.0	411	12	B86567	B86567 PC111-24C1
67	19	1.0	433	10	BG077137	BG077137 H3011C04-
68	19	1.0	450	12	A2143510	A2143510 SP_0018_A
69	19	1.0	465	12	AQ013306	AQ013306 RPT11-24
70	19	1.0	470	12	A2086245	A2086245 RPT1-23-4
71	19	1.0	484	10	BF994356	BF994356 MR0-GN017
72	19	1.0	489	10	BF775254	BF775254 285365 MA
73	19	1.0	497	12	A2763586	A2763586 1M0559B08
74	19	1.0	501	12	BH217548	BH217548 1006056C0
75	19	1.0	527	12	A2759820	A2759820 1M0552L24
76	19	1.0	531	10	BM334118	BM334118 MEST132-H
77	19	1.0	541	10	BI974963	BI974963 sal15c11.
78	19	1.0	543	10	BE460169	BE460169 EST415461
79	19	1.0	566	10	BI705962	BI705962 fP98F03.Y
80	19	1.0	581	10	BE459685	BE459685 EST41497.Y
81	19	1.0	581	12	AO788778	AO788778 HS_3187_B
82	19	1.0	593	12	BE458253	BE458253 EST413593
83	19	1.0	608	12	BH259200	BH259200 CH230-29B
84	19	1.0	613	12	AO385601	AO385601 RPT11-13
85	19	1.0	626	10	BM490620	BM490620 pP2p2n.pK0
86	19	1.0	635	12	A2409586	A2409586 1M0181615
87	19	1.0	660	10	BH423208	BH423208 BOCMF12TF
88	19	1.0	665	12	BH062477	BH062477 RPT1-24-3
89	19	1.0	677	9	AI588624	AI588624 fb97d10.y
90	19	1.0	677	10	BI661446	BI661446 603304741

91	19	1.0	681	10	BI099954	602885082	164	18	0.9	403	12	AQ253985	AQ253985 HS-2225-A
92	19	1.0	698	12	AZ908847	RPCI-24-1	165	18	0.9	406	9	AM168430	AM168430 x183d04.x
93	19	1.0	721	10	BF581163	602100340	166	18	0.9	406	9	BF303196	BF303196 f184809.x
94	19	1.0	725	12	AZ629815	1M0483306	167	18	0.9	407	10	BI578309	BI578309 RE72162.5
95	19	1.0	752	12	BH244978	A028A71TR	168	18	0.9	408	10	BI485888	BI485888 RK69241.5
96	19	1.0	758	10	BG963476	602831137	169	18	0.9	411	9	AV729734	AV729734 AV729734
97	19	1.0	776	10	BI102190	602887975	170	18	0.9	415	10	BI484107	BI484107 RB67141.5
98	19	1.0	785	12	BH583112	BH583112 BOGEN26TF	171	18	0.9	421	9	AU208871	AU208871 AU208871
99	19	1.0	801	12	BH482780	BH482780 BOHFEP6TF	172	18	0.9	423	9	AI167181	AI167181 OKx28E04.x
100	19	1.0	805	12	BH511534	BH511534 BCGBD36TF	173	18	0.9	423	12	BH00034	BH00034 CIT-HSP-204
101	19	1.0	809	12	BF236718	BF236718 602028013	174	18	0.9	427	10	BI005041	BI005041 PM3-HN007
102	19	1.0	859	12	BH111314	BH111314 RPCI-24-3	175	18	0.9	430	10	AM477317	AM477317 ga45c11.y
103	19	1.0	876	10	BF263156	BF263156 HY-CEA000	176	18	0.9	431	12	AQ584924	AQ584924 RPCI-11-4
104	19	1.0	906	12	AZ689698	ENTK085TF	177	18	0.9	436	10	C82907	C82907 C82907 rabd
105	19	1.0	910	12	AZ550500	ENTEC19TF	178	18	0.9	436	10	C83763	C83763 C83763 rabd
106	19	1.0	924	10	BF584742	BF584742 602098492	179	18	0.9	437	10	W35196	W35196 zc70d06.r1
107	19	1.0	940	10	BG298341	BG298341 602397127	180	18	0.9	438	9	AI475337	AI475337 f182a04.x
108	19	1.0	1038	12	CNS01HMM	AL144659 Anopheles	181	18	0.9	438	10	BM440956	BM440956 EBEd02-SQ
109	19	1.0	2199	11	AK019222	AK019222 Mus muscu	182	18	0.9	438	10	R73286	R73286 yJ92d01.r1
110	19	0.9	117	10	T63516	T63516 yc07a03.s1	183	18	0.9	438	12	AZ475157	AZ475157 IM0293B11
111	18	0.9	124	12	BH417338	BH417338 1007054D0	184	18	0.9	439	10	C06769	C06769 C06769 Rat
112	18	0.9	174	12	AO269185	RPCI11-75	185	18	0.9	441	12	AO842111	AO842111 T134A07.S
113	18	0.9	183	10	BG968819	BG968819 602836284	186	18	0.9	442	10	BF222120	BF222120 7p42f03.x
114	18	0.9	187	9	AA7711964	AA7711964 v029c04.r	187	18	0.9	442	12	AQ265185	AQ265185 CTBI-EL-
115	18	0.9	207	12	B96425	B96425 F19B14TR.IG	188	18	0.9	449	12	AQ695311	AQ695311 HS-2142-A
116	18	0.9	208	12	AM859435	AM859435 MR1-CT035	189	18	0.9	450	10	BE497833	BE497833 WHE0957.F
117	18	0.9	208	12	BH206244	BH206244 Sm1-49H7.	190	18	0.9	452	9	AA293412	AA293412 zt53f09.r
118	18	0.9	217	9	AV334982	AV334982 AV334982	191	18	0.9	454	9	AA468215	AA468215 hc34c01.x
119	18	0.9	234	9	AM859384	AM859384 MR1-CT035	192	18	0.9	459	9	AW770104	AW770104 h183g01.x
120	18	0.9	253	9	AA356650	AA356650 EST65492	193	18	0.9	460	10	W95864	W95864 ze08a01.s1
121	18	0.9	257	9	AA322319	AA322319 EST24941	194	18	0.9	461	10	BI797071	BI797071 H063H02.E
122	18	0.9	271	10	NS59046	NS59046 yv60c12.r1	195	18	0.9	461	12	AZ811979	AZ811979 2M0087D14
123	18	0.9	272	10	AG940864	AG940864 ax07h09.y	196	18	0.9	462	9	AI588336	AI588336 fbd9e05.x
124	18	0.9	279	12	AA934680	AA934680 o080b12.s	197	18	0.9	463	10	BE441391	BE441391 925005C08
125	18	0.9	279	12	AO068440	AO068440 HS-2243-B	198	18	0.9	464	10	BI164963	BI164963 RE04744.5
126	18	0.9	283	10	BF938266	BF938266 fm74e06.y	199	18	0.9	469	9	AI811634	AI811634 tw46a10.x
127	18	0.9	286	10	T93749	T93749 ye10d02.r1	200	18	0.9	469	9	AI840859	AI840859 UI-M-AH0-
128	18	0.9	288	10	BI005131	BI005131 PM3-HN007	201	18	0.9	473	10	BI137007	BI137007 F078P35Y
129	18	0.9	289	9	BB228355	BB228355 BB228355	202	18	0.9	473	10	BI124998	BI124998 BI124998
130	18	0.9	302	9	BE074170	RC4-BR056	203	18	0.9	473	10	BE442227	BE442227 925014H07
131	18	0.9	302	9	AI945138	bs09d08.y	204	18	0.9	474	9	AI994065	AI994065 701498437
132	18	0.9	311	10	BB164426	BB164426 BB164426	205	18	0.9	475	9	AI944429	AI944429 bs01c09.y
133	18	0.9	311	10	BM238980	BM238980 K0528G08-	206	18	0.9	477	10	BG639388	BG639388 LP07744.3
134	18	0.9	323	12	AQ382777	AQ382777 RPCI11-13	207	18	0.9	482	10	R95815	R95815 yq50e07.r1
135	18	0.9	325	9	AM565951	s188h03.y	208	18	0.9	483	10	BI592216	BI592216 RH09855.5
136	18	0.9	325	10	BG940862	ax07h09.y	209	18	0.9	486	9	AA254066	AA254066 va10f05.r
137	18	0.9	332	9	AV689732	AV689732 AV689732	210	18	0.9	486	9	AA589206	AA589206 vx24e01.r
138	18	0.9	334	10	BI473730	BI473730 fp41f04.y	211	18	0.9	486	10	BM239585	BM239585 K0537A12-
139	18	0.9	347	9	AA658202	AA658202 nu20D06.s	212	18	0.9	487	9	AA395358	AA395358 27135.Lam
140	18	0.9	355	12	AZ904101	AZ904101 RPCI-24-1	213	18	0.9	490	10	BF623829	BF623829 HVM5EA000
141	18	0.9	358	9	AM258934	AM258934 um75a08.y	214	18	0.9	491	9	AI654525	AI654525 tg91g06.x
142	18	0.9	360	9	AI301700	AI301700 qm36908.x	215	18	0.9	491	10	BE857967	BE857967 7f73d04.x
143	18	0.9	367	10	BF43122	BF43122 260495.MA	216	18	0.9	494	9	AA576752	AA576752 nm81g08.s
144	18	0.9	368	9	AM273218	AM273218 x335e08.x	217	18	0.9	496	10	BG849419	BG849419 1024025C0
145	18	0.9	369	10	BI228681	BI228681 RE26305.5	218	18	0.9	496	10	BI366957	BI366957 RE52612.5
146	18	0.9	369	12	AZ503052	AZ503052 IM0342111	219	18	0.9	497	12	AQ941496	AQ941496 Sheared.D
147	18	0.9	370	9	AM275317	AM275317 xv80h12.x	220	18	0.9	500	10	BI482175	BI482175 BE64833.5
148	18	0.9	372	12	BH232336	BH232336 100616780	221	18	0.9	500	12	AZ021229	AZ021229 RPCI-23-3
149	18	0.9	374	9	AA504438	AA504438 aa59g11.r	222	18	0.9	504	9	AM198586	AM198586 p115.pK0
150	18	0.9	376	10	C69781	C69781 C69781.YUJ1	223	18	0.9	505	9	AI742888	AI742888 wq47e02.x
151	18	0.9	377	9	AV729733	AV729733 AV729733	224	18	0.9	505	10	BI239057	BI239057 RE35447.5
152	18	0.9	383	9	AV188421	AV188421 AV188421	225	18	0.9	505	10	BF654492	BF654492 278653.MA
153	18	0.9	383	12	AZ904732	AZ904732 RPCI-24-1	226	18	0.9	507	12	AZ846966	AZ846966 2M0147C23
154	18	0.9	384	9	AA164846	AA164846 zp02b11.r	227	18	0.9	508	12	AO700726	AO700726 HS-5321.A
155	18	0.9	390	12	AO314227	AO314227 RPCI11-96	228	18	0.9	510	10	BI235507	BI235507 RE31359.5
156	18	0.9	393	9	AA536345	AA536345 LD16910.5	229	18	0.9	511	10	BI136700	BI136700 F072P19Y
157	18	0.9	393	10	BE834619	BE834619 MR1-FN000	230	18	0.9	512	9	AI953689	AI953689 wq47c01.x
158	18	0.9	396	12	AO511746	AO511746 HS-f119.A	231	18	0.9	513	9	BI367740	BI367740 RE53572.5
159	18	0.9	399	12	AO152510	AO152510 HS-3110.B	232	18	0.9	517	10	BI004814	BI004814 BI004814
160	18	0.9	400	9	AI394888	AI394888 MA001297.	233	18	0.9	517	10	BI112673	BI112673 BI112673
161	18	0.9	401	9	AI385055	AI385055 MA002302.	234	18	0.9	518	10	BI169087	BI169087 RE10064.5
162	18	0.9	403	9	AA025915	AA025915 ze91c03.r	235	18	0.9	518	10	BI804887	BI804887 S001F07.S
163	18	0.9	403	10	Z34597	Z34597 ATTS3357.Ve	236	18	0.9	518	10	BM301346	BM301346 MCR041F08

C 237	18	0.9	520	12	A2172702	SP_0122_B	310	18	0.9	615	9	AU034816	AU034816
C 238	18	0.9	521	10	B1368754	RE54779.5	311	18	0.9	616	12	A2947246	A2947246
C 239	18	0.9	524	12	A0017893	CIT-HSP_-2	312	18	0.9	617	10	B1212951	B1212951
C 240	18	0.9	525	10	B6672082	7a51c01.x	313	18	0.9	617	12	A2800372	A2800372
C 241	18	0.9	527	12	B64374	CIT-HSP_-202	314	18	0.9	618	9	AV401322	AV401322
C 242	18	0.9	529	10	B1227672	RE24781.5	315	18	0.9	618	9	BD498288	BD498288
C 243	18	0.9	529	10	BM116635	BM116635 L0839B04	316	18	0.9	618	10	B1364444	B1364444
C 244	18	0.9	530	12	AM055074	AM055074 WY9C09.x	317	18	0.9	619	10	B1216624	B1216624
C 245	18	0.9	531	10	BG150430	BG150430 7k01a02.x	318	18	0.9	619	10	B1374865	B1374865
C 246	18	0.9	532	9	AT587545	AT587545 t552b07.x	319	18	0.9	619	12	A2334737	A2334737
C 247	18	0.9	532	9	AT587545	AT587545 t552b07.x	320	18	0.9	620	10	B1171092	B1171092
C 248	18	0.9	532	12	AT501525	AT501525 1M03401122	321	18	0.9	621	10	B1364570	B1364570
C 249	18	0.9	534	12	BE590370	BE590370 nab32901.	322	18	0.9	622	10	AL577457	AL577457
C 250	18	0.9	534	12	B85229	B85229 RPTC11-30M7	323	18	0.9	622	12	AQ449509	AQ449509
C 251	18	0.9	535	9	AT483218	AT483218 EST242695	324	18	0.9	623	10	BM159288	BM159288
C 252	18	0.9	535	9	AT483218	AT483218 t08f03.x	325	18	0.9	624	9	BB630750	BB630750
C 253	18	0.9	537	9	AT608110	AT608110 vA10f05.y	326	18	0.9	627	10	B1162430	B1162430
C 254	18	0.9	540	10	B1375957	B1375957 RE63866.5	327	18	0.9	629	10	BF863372	BF863372
C 255	18	0.9	540	12	AQ753968	HS_5398_A	328	18	0.9	630	12	BH056545	BH056545
C 256	18	0.9	542	10	BJ116355	BJ116355 BJ116355	329	18	0.9	631	9	AT357593	AT357593
C 257	18	0.9	542	10	BF196828	BF196828 7n07d04.x	330	18	0.9	631	10	BJ006434	BJ006434
C 258	18	0.9	545	10	BE429899	BE429899 TAS004.H1	331	18	0.9	632	9	AT774507	AT774507
C 259	18	0.9	547	9	AT675148	AT675148 wC09h11.x	332	18	0.9	633	9	AA201744	AA201744
C 260	18	0.9	547	10	BM407754	BM407754 EST582081	333	18	0.9	634	9	AM696903	AM696903
C 261	18	0.9	550	12	AQ281037	AQ281037 RPTC11-81	334	18	0.9	634	9	BB208421	BB208421
C 262	18	0.9	551	10	BI486090	BI486090 RE69463.5	335	18	0.9	639	10	B1214545	B1214545
C 263	18	0.9	551	10	BE575706	BE575706 211961.MA	336	18	0.9	641	10	BJ108064	BJ108064
C 264	18	0.9	551	12	A2169690	SP_0114_A	337	18	0.9	641	10	BJ114124	BJ114124
C 265	18	0.9	553	10	BI236582	RE32653.5	338	18	0.9	644	10	BE273985	BE273985
C 266	18	0.9	555	10	BI228974	RE26650.5	339	18	0.9	644	12	BH442086	BH442086
C 267	18	0.9	557	12	B98152	B98152 F26D24TFB I	340	18	0.9	645	12	A2804232	A2804232
C 268	18	0.9	558	9	AT656747	AT656747 t553n08.x	341	18	0.9	649	9	BB160674	BB160674
C 269	18	0.9	558	10	BM108987	BM108987 EST556523	342	18	0.9	651	9	BB159570	BB159570
C 270	18	0.9	559	10	BJ111568	BJ111568 BJ111568	343	18	0.9	653	9	AM002316	AM002316
C 271	18	0.9	562	9	AU034779	AU034779 AU034779	344	18	0.9	653	10	BI215969	BI215969
C 272	18	0.9	562	10	BI375009	BI375009 RE62715.5	345	18	0.9	654	10	BI356777	BI356777
C 273	18	0.9	563	10	BE442362	BE442362 925019C05	346	18	0.9	654	12	BH091880	BH091880
C 274	18	0.9	566	9	AV610755	AV610755 AV610755	347	18	0.9	655	10	BM163651	BM163651
C 275	18	0.9	567	12	A2093968	RPTC_-23-4	348	18	0.9	655	12	BM173717	BM173717
C 276	18	0.9	568	10	BE594413	BE594413 RE38396.5	349	18	0.9	659	10	BI238697	BI238697
C 277	18	0.9	572	9	AA990731	AA990731 P11_33_B0	350	18	0.9	660	10	BI374878	BI374878
C 278	18	0.9	573	10	BI237259	BI237259 RE33451.5	351	18	0.9	660	10	BM160463	BM160463
C 279	18	0.9	573	10	BE554390	ur42a05.y	352	18	0.9	661	10	BM231189	BM231189
C 280	18	0.9	575	12	BH405516	BH405516 GM_15B001	353	18	0.9	662	9	AV880405	AV880405
C 281	18	0.9	575	12	BH405516	BH405516 GM_15B001	354	18	0.9	665	12	BH615359	BH615359
C 282	18	0.9	579	10	BE506173	BE506173 AT06574.5	355	18	0.9	666	10	BI228862	BI228862
C 283	18	0.9	580	9	AA521571	AA521571 v115602.T	356	18	0.9	670	10	BI282068	BI282068
C 284	18	0.9	581	10	BI173592	BI173592 RE16542.5	357	18	0.9	670	10	BI483728	BI483728
C 285	18	0.9	581	10	BI237311	BI237311 RE33523.5	358	18	0.9	670	10	BE502379	BE502379
C 286	18	0.9	584	10	BM136167	BM136167 WHEZ606.A	359	18	0.9	671	10	BM171161	BM171161
C 287	18	0.9	584	10	W80408	W80408 zH49f05.s1	360	18	0.9	672	10	BI485537	BI485537
C 288	18	0.9	585	10	BG523572	33-28_Ste	361	18	0.9	675	10	BG633677	BG633677
C 289	18	0.9	585	10	BM145014	TCAP1D14	362	18	0.9	676	10	BI228209	BI228209
C 290	18	0.9	587	10	BM163467	BM163467 EST565990	363	18	0.9	677	10	BI579195	BI579195
C 291	18	0.9	588	12	A2212241	A2212241 Sheared.D	364	18	0.9	679	10	BM170305	BM170305
C 292	18	0.9	588	12	A2801390	A2801390 2M0059P13	365	18	0.9	679	12	AG059706	AG059706
C 293	18	0.9	589	9	AT1719073	as65e08.x	366	18	0.9	679	12	AG072134	AG072134
C 294	18	0.9	589	9	AT1971249	AT1971249 wr27a10.x	367	18	0.9	681	9	AT1774508	AT1774508
C 295	18	0.9	589	12	A2856098	A2856098 2M0160H19	368	18	0.9	684	10	BG635702	BG635702
C 296	18	0.9	590	9	AT1760647	AT1760647 w166904.x	369	18	0.9	686	9	AM969044	AM969044
C 297	18	0.9	592	10	BI232447	BI232447 RE28515.5	370	18	0.9	687	10	BJ004084	BJ004084
C 298	18	0.9	596	9	AT1928200	AT1928200 w095c11.x	371	18	0.9	688	10	BI231283	BI231283
C 299	18	0.9	596	10	BR491199	BR491199 AT27886.5	372	18	0.9	688	10	BM009398	BM009398
C 300	18	0.9	598	9	AA390638	AA390638 LD08837.5	373	18	0.9	691	9	BE208978	BE208978
C 301	18	0.9	602	10	BI376051	BI376051 RE63975.5	374	18	0.9	692	12	AG114407	AG114407
C 302	18	0.9	605	10	BG180550	BG180550 602331483	375	18	0.9	693	9	AU101111	AU101111
C 303	18	0.9	609	9	AV385599	AV385599 AV385599	376	18	0.9	695	10	BI236610	BI236610
C 304	18	0.9	609	12	A2470485	A2470485 1M0284C01	377	18	0.9	698	10	BI368144	BI368144
C 305	18	0.9	610	10	BI362389	BI362389 RE46933.5	378	18	0.9	699	10	BM407630	BM407630
C 306	18	0.9	611	10	BJ025129	BJ025129 BJ025129	379	18	0.9	700	9	AL508089	AL508089
C 307	18	0.9	611	10	BE423488	BE423488 WHE0064_G	380	18	0.9	704	10	BG491603	BG491603
C 308	18	0.9	611	12	AQ324245	AQ324245 mgxB0017F	381	18	0.9	704	10	BI231108	BI231108
C 309	18	0.9	613	12	AQ681648	AQ681648 HS_2161_A	382	18	0.9	705	12	BH058352	BH058352

AU034816	AU034816
AZ494746	2M0209F23
B112951	RE18231.5
A2800372	2M0058E06
AV401322	AV401322
BD498288	BD498288
B1364444	RE49548.5
B1216624	RE23825.5
B1374865	RE62550.5
A2334737	1M0064C14
BI171092	RE12768.5
BI664570	RE49725.5
AL577457	AL577457
AQ449509	500001D01
BM159288	EST561811
BB630750	BB630750
BI162430	RE01750.5
BF863372	963043B11
BH056545	RPTC_-24-3
AT357593	QY14h07.x
BJ006434	BJ006434
AT774507	EST255607
AA201744	LD04957.5
AM696903	NE110C03S
BB208421	BB208421
BI214545	RE20380.5
BJ108064	BJ108064
BJ114124	BJ114124
BE273985	60110446F
BH442086	BOHMA44TF
A2804232	2M0065J23
BB160674	BB160674
BB159570	BB159570
AM002316	vW45608.x
BI215969	RE22750.5
BI356777	RE42983.5
BH091880	RPTC_-24-3
BM163651	EST566174
BM173717	A407/F_RF
BI238697	RE35214.5
BI374878	RE62563.5
BM160463	EST562986
BM231189	K0306D11-
AV880405	AV880405
BH615359	BMAAC302A
BI228862	RE25480.5
BI238238	RE34642.5
BF402379	RE66652.5
BE502379	AT18080.5
BM171161	EST573684
BI485537	RE68887.5
BG633677	GM06560.3
BI228209	RE25674.5
BI579195	RE73184.5
BM170305	EST572828
AG059706	Pan t10g1
AG072134	Pan t10g1
AT1774508	EST255608
BG635702	SD13165.5
AM969044	EST381121
BJ004084	BJ004084
BI231283	RE20730.5
BM009398	603629881
BE208978	GF-FV-PAD
AG114407	Pan t10g1
AU101111	AU101111
BI236610	RE32682.5
BI368144	RE54063.5
BM407630	EST581957
AL508089	AL508089
BG491603	602535930
BI231108	RE19301.5
BH058352	RPTC_-24-3

383	18	0.9	706	10	B1559365	603253036	C 456	18	0.9	995	9	AL572762	AL572762
384	18	0.9	709	10	B1484189	RE67248.5	C 457	18	0.9	1024	10	BM423728	BM423728
C 385	18	0.9	713	10	B1522087	603081815	C 458	18	0.9	1010	10	BG822861	BG822861
C 386	18	0.9	714	10	BE225469	MD0377 Me	C 459	18	0.9	1036	12	CNS01G4N	CNS01G4N
387	18	0.9	714	12	BH118334	RPCI-24-2	C 460	18	0.9	1040	12	CNS02NAX	CNS02NAX
388	18	0.9	716	10	B1215435	RE21836.5	C 461	18	0.9	1042	9	AL517775	AL517775
389	18	0.9	716	10	BM239189	K0530D07-	C 462	18	0.9	1049	10	BM457075	BM457075
390	18	0.9	717	9	AV401272	AV401272	C 463	18	0.9	1050	9	AL582276	AL582276
391	18	0.9	717	9	BE071793	RCO-BT052	C 464	18	0.9	1089	10	BG787895	BG787895
C 392	18	0.9	719	12	BH013921	PCGREF0TH	C 465	18	0.9	1100	12	CNS055AM	CNS055AM
C 393	18	0.9	719	12	AQ361498	meB0004K	C 466	18	0.9	1415	10	BG845565	BG845565
C 394	18	0.9	724	9	AW515087	xu91b07.x	C 467	18	0.9	114	12	AZ402379	AZ402379
C 395	18	0.9	733	12	BH570007	BOHNNR0TR	C 468	18	0.9	116	12	AQ630915	AQ630915
C 396	18	0.9	738	10	BF506094	AT08477.5	C 469	18	0.9	1330	10	B11996618	B11996618
C 397	18	0.9	743	10	BM162957	EST565480	C 470	18	0.9	1332	10	BM110457	BM110457
C 398	18	0.9	743	10	BF037425	601461002	C 471	18	0.9	137	10	BF014246	BF014246
399	18	0.9	750	10	BE876638	601488201	C 472	18	0.9	153	10	BF329119	BF329119
C 400	18	0.9	754	9	AL526524	AL526524	C 473	18	0.9	163	9	AU213065	AU213065
C 401	18	0.9	755	10	BE875180	601488667	C 474	18	0.9	166	10	BE757712	BE757712
C 402	18	0.9	763	12	BH492670	BOHGC30TR	C 475	18	0.9	167	10	BG868566	BG868566
C 403	18	0.9	770	12	BH101159	RPCI-24-3	C 476	18	0.9	173	9	AW796288	AW796288
C 404	18	0.9	777	10	AU000571	AU000571	C 477	18	0.9	182	10	BE841160	BE841160
C 405	18	0.9	778	10	BM010510	603631475	C 478	18	0.9	186	9	AA502132	AA502132
C 406	18	0.9	779	9	AL623065	AENTAD21	C 479	18	0.9	189	9	AA806651	AA806651
C 407	18	0.9	780	10	BM159621	EST562144	C 480	18	0.9	191	9	AA422649	AA422649
C 408	18	0.9	783	9	AL563059	AL563059	C 481	18	0.9	191	9	AA619734	AA619734
C 409	18	0.9	788	9	AV400991	AV400991	C 482	18	0.9	197	9	AA681308	AA681308
C 410	18	0.9	791	10	AA567948	HL02035.5	C 483	18	0.9	197	9	AW659556	AW659556
C 411	18	0.9	791	10	BG760386	602716807	C 484	18	0.9	200	9	AL850166	AL850166
C 412	18	0.9	793	12	BH471988	BGFT36TR	C 485	18	0.9	200	10	BF774407	BF774407
C 413	18	0.9	799	10	BG389916	602415459	C 486	18	0.9	203	10	F15156	F15156
C 414	18	0.9	799	12	BH490388	BOHBI13TR	C 487	18	0.9	206	10	BE821671	BE821671
C 415	18	0.9	802	10	BF968244	602269611	C 488	18	0.9	211	10	BG017720	BG017720
C 416	18	0.9	815	10	BG780345	SEAUWC000	C 489	18	0.9	213	10	T48934	T48934
C 417	18	0.9	823	9	AV401129	AV401129	C 490	18	0.9	215	9	BM145109	BM145109
C 418	18	0.9	835	9	AL572467	AL572467	C 491	18	0.9	217	9	AI968081	AI968081
C 419	18	0.9	841	9	AL579300	AG137000	C 492	18	0.9	218	10	BF546288	BF546288
C 420	18	0.9	841	12	AG137000	PH Trogl	C 493	18	0.9	220	9	AV335016	AV335016
C 421	18	0.9	844	12	BH023317	GH MB000	C 494	18	0.9	224	9	AA566015	AA566015
C 422	18	0.9	844	12	BH511647	BGCF71TR	C 495	18	0.9	226	10	B1316911	B1316911
C 423	18	0.9	846	10	BE972796	601652028	C 496	18	0.9	227	10	B1435280	B1435280
C 424	18	0.9	849	10	BF527018	602070612	C 497	18	0.9	234	10	BE484532	BE484532
C 425	18	0.9	849	12	BH583147	BGGM135TR	C 498	18	0.9	235	10	H21312	H21312
C 426	18	0.9	854	12	AQ450136	500008D02	C 499	18	0.9	236	10	BE725065	BE725065
C 427	18	0.9	856	10	BE797809	601586425	C 500	18	0.9	240	9	AU074507	AU074507
C 428	18	0.9	858	10	BG623369	602586413	C 501	18	0.9	243	12	AZ258446	AZ258446
C 429	18	0.9	867	9	AL574257	AL574257	C 502	18	0.9	244	9	AV273323	AV273323
C 430	18	0.9	869	10	BG570663	602590965	C 503	18	0.9	244	10	BE527808	BE527808
C 431	18	0.9	871	10	BM006012	603613748	C 504	18	0.9	246	9	AA570408	AA570408
C 432	18	0.9	872	10	B1522040	603081736	C 505	18	0.9	247	9	AW706163	AW706163
C 433	18	0.9	874	9	AL571869	AL571869	C 506	18	0.9	249	9	AI556653	AI556653
C 434	18	0.9	877	10	BG476325	602521818	C 507	18	0.9	256	9	BB579437	BB579437
C 435	18	0.9	881	12	BH160577	EMTQ11TR	C 508	18	0.9	256	9	BB597624	BB597624
C 436	18	0.9	884	10	BG330475	602430147	C 509	18	0.9	257	9	AV163047	AV163047
C 437	18	0.9	902	9	AL550497	AL550497	C 510	18	0.9	261	9	BB010222	BB010222
C 438	18	0.9	906	9	AL580498	AL580498	C 511	18	0.9	262	10	B1015339	B1015339
C 439	18	0.9	908	12	BM010512	603631477	C 512	18	0.9	264	9	BB748332	BB748332
C 440	18	0.9	908	12	CNS05B19	AL339238	C 513	18	0.9	266	10	BF220401	BF220401
C 441	18	0.9	909	10	BG423508	602449370	C 514	18	0.9	268	9	BA466539	BA466539
C 442	18	0.9	910	12	CNS036T1	AL230446	C 515	18	0.9	268	10	B1922332	B1922332
C 443	18	0.9	911	12	CNS06J3K	Tetraodon	C 516	18	0.9	268	12	CNS00VD6	CNS00VD6
C 444	18	0.9	912	10	B1078938	602873184	C 517	18	0.9	272	9	AV300066	AV300066
C 445	18	0.9	912	10	BE743286	601574885	C 518	18	0.9	274	9	BB195297	BB195297
C 446	18	0.9	929	10	BF791469	602251511	C 519	18	0.9	277	9	BB351790	BB351790
C 447	18	0.9	945	10	B1089149	602854004	C 520	18	0.9	277	9	BB722075	BB722075
C 448	18	0.9	945	10	BE617956	601442184	C 521	18	0.9	286	9	AW568359	AW568359
C 449	18	0.9	947	9	AL579071	AL579071	C 522	18	0.9	286	10	BF186611	BF186611
C 450	18	0.9	949	9	BE213916	BE213916	C 523	18	0.9	289	10	BG239086	BG239086
C 451	18	0.9	962	12	CNS07DB3	AL440245	C 524	18	0.9	289	12	B06423	B06423
C 452	18	0.9	970	12	CNS04AMW	AL282425	C 525	18	0.9	290	9	BE180963	BE180963
C 453	18	0.9	975	9	AL572942	AL572942	C 526	18	0.9	290	10	BE477329	BE477329
C 454	18	0.9	978	10	BM424224	AGENCOURT	C 527	18	0.9	291	9	BB194398	BB194398
C 455	18	0.9	978	10	BE740965	601595328	C 528	18	0.9	294	9	AV293940	AV293940

529	17	0.9	295	10	BE824216	BE824216 GM700023A	602	17	0.9	370	9	AU204179	AU204179
530	17	0.9	296	10	BF391767	BF391767 UI-R-CA1-	603	17	0.9	371	12	A2127658	A2127658 OSUNB009
531	17	0.9	299	9	AM334777	AM334777 S378 ACS	604	17	0.9	370	10	BE802196	BE802196 sr25f05.y
532	17	0.9	300	9	AA229159	AA229159 nc45B11.s	605	17	0.9	375	10	AV198561	AV198561 AV198561
533	17	0.9	300	9	BA400928	BA400928	606	17	0.9	375	10	CA3941	CA3941 CA3941 yuji
534	17	0.9	302	9	AV167272	AV167272	607	17	0.9	376	12	AQ045322	AQ045322 RPTC111-36
535	17	0.9	302	9	BB381934	BB381934	608	17	0.9	377	9	AM788524	AM788524 C00516-F
536	17	0.9	303	10	BE317699	BE317699 NF053F03L	609	17	0.9	377	9	BB804538	BB804538
537	17	0.9	304	10	BF964473	BF964473 RCL-NN023	610	17	0.9	377	10	BG159315	BG159315 OV2.6.C01
538	17	0.9	304	10	BF054642	BF054642 EST439872	611	17	0.9	378	10	BG159315	BG159315
539	17	0.9	307	9	BB838179	BB838179	612	17	0.9	378	10	B1359863	B1359863
540	17	0.9	308	10	BF958519	BF958519 OV2-NN004	613	17	0.9	378	10	CA8533	CA8533
541	17	0.9	310	9	AA793720	AA793720 vr95g10.r	614	17	0.9	378	12	A2493775	A2493775 yuji
542	17	0.9	311	9	AA947333	AA947333 Oq49508.s	615	17	0.9	379	12	A0010574	A0010574
543	17	0.9	311	9	B1033010	B1033010 PMO-MT042	616	17	0.9	380	10	CA7511	CA7511
544	17	0.9	311	10	BT033021	BT033021 PMO-MT042	617	17	0.9	381	12	A0061280	A0061280 CIT-HSP-2
545	17	0.9	312	9	AM347101	AM347101 30235.MAR	618	17	0.9	382	9	A1923125	A1923125
546	17	0.9	315	10	UB5028	UB5028 DM085028.Dr	619	17	0.9	382	12	AQ959102	AQ959102 LERREH30TR
547	17	0.9	319	10	B1316939	B1316939 saf82e07.	620	17	0.9	383	9	AA928151	AA928151 ONR6E01.s
548	17	0.9	320	9	AM659554	AM659554 96860.MAR	621	17	0.9	383	12	AQ967181	AQ967181 LERIO14TF
549	17	0.9	320	10	H11416	H11416 yml13q05.r1	622	17	0.9	384	12	AM543420	AM543420 C0169E10-
550	17	0.9	321	10	BE940989	BE940989 EST420568	623	17	0.9	384	12	A0804601	A0804601 HS-2263_B
551	17	0.9	322	9	AM484678	AM484678 62094.MAR	624	17	0.9	385	10	T47128	T47128 yb52901.r1
552	17	0.9	323	9	BB100939	BB100939	625	17	0.9	386	9	BB804917	BB804917
553	17	0.9	325	10	BG158832	BG158832 RH122.43_	626	17	0.9	386	12	BH126745	BH126745 BARC-Salt
554	17	0.9	330	9	AA366781	AA366781 EST77771	627	17	0.9	388	10	BG644749	BG644749 EST506368
555	17	0.9	330	10	B1372506	B1372506 RES9481.5	628	17	0.9	388	12	BH016030	BH016030 TDCGH56TH
556	17	0.9	330	10	BT398532	BT398532 952002E01	629	17	0.9	389	9	AM987158	AM987158 w15b05.x
557	17	0.9	334	12	CNS01RZA	AL164575 Tetradon	630	17	0.9	389	12	A2868866	A2868866 2M0180E13
558	17	0.9	335	12	A2849782	A2849782 2M0151B24	631	17	0.9	390	10	B1926856	B1926856 EST546745
559	17	0.9	336	12	A2924995	A2924995 4910.ez31	632	17	0.9	390	12	A2752618	A2752618 RPTC-24-1
560	17	0.9	337	9	AM217308	AM217308 EST296027	633	17	0.9	391	9	AM450960	AM450960 UI-H-B13-
561	17	0.9	338	9	AM569761	AM569761 s181b01.y	634	17	0.9	392	9	AL635835	AL635835
562	17	0.9	339	10	C20724	C20724 HMG800475	635	17	0.9	392	12	BH098871	BH098871 RPTC-24-3
563	17	0.9	340	9	A1363761	A1363761 qv79802.x	636	17	0.9	393	12	B30710	B30710 HS-1003-A1-
564	17	0.9	341	10	F20510	F20510 HSPD04860.H	637	17	0.9	394	10	BG134036	BG134036 EST466928
565	17	0.9	342	9	AU176520	AU176520	638	17	0.9	397	9	AA558145	AA558145 n126d10.s
566	17	0.9	344	10	BM442866	BM442866 EBR002.SQ	639	17	0.9	398	9	A1273733	A1273733
567	17	0.9	345	10	B1347479	B1347479	640	17	0.9	398	10	BG927311	BG927311 HNC64-1-C
568	17	0.9	346	12	A0756347	A0756347 HS-5210_B	641	17	0.9	399	9	AA386929	AA386929
569	17	0.9	346	9	AV197804	AV197804	642	17	0.9	399	10	N23811	N23811
570	17	0.9	346	9	AA400475	AA400475 zu64e10.s	643	17	0.9	399	12	AQ121279	AQ121279 HS-3075_A
571	17	0.9	346	10	T05409	T05409 EST03298.Fe	644	17	0.9	400	10	BE571388	BE571388
572	17	0.9	346	10	BF609597	BF609597 NXSI_047_	645	17	0.9	401	9	AM086060	AM086060 xc77g03.x
573	17	0.9	347	12	N50424	N50424 YYR80808.s1	646	17	0.9	402	10	BF526719	BF526719
574	17	0.9	347	12	A2776537	A2776537 2M0010C18	647	17	0.9	403	9	AA105576	AA105576
575	17	0.9	348	10	B1934423	B1934423 EST554312	648	17	0.9	403	12	BH140250	BH140250
576	17	0.9	348	12	BH004658	BH004658 BMAC08G0	649	17	0.9	404	10	W97823	W97823
577	17	0.9	350	9	AU112203	AU112203	650	17	0.9	405	9	AM008231	AM008231
578	17	0.9	351	9	A1825450	A1825450 wb73b01.x	651	17	0.9	405	10	BG137180	BG137180 EST477622
579	17	0.9	351	9	AU112508	AU112508	652	17	0.9	406	9	BE220007	BE220007 hv65N05.x
580	17	0.9	351	10	BE756537	BE756537 210705.MA	653	17	0.9	407	9	A1821281	A1821281
581	17	0.9	351	10	BE850184	BE850184 uw13h11.y	654	17	0.9	407	9	A1945739	A1945739
582	17	0.9	352	10	BF044171	BF044171 BP250012B	655	17	0.9	407	10	BG012541	BG012541 ILS-GN03
583	17	0.9	355	9	AV011169	AV011169	656	17	0.9	408	10	BF774413	BF774413 284079.MA
584	17	0.9	355	10	B1972374	B1972374 saq90f01.	657	17	0.9	408	12	B72961	B72961 RPTC11-10M1
585	17	0.9	357	9	AM085070	AM085070 xe05C01.x	658	17	0.9	409	10	BF939304	BF939304 nad6b04.
586	17	0.9	358	9	AM614412	AM614412 hg90g10.x	659	17	0.9	409	10	B1112593	B1112593
587	17	0.9	360	9	AU098178	AU098178	660	17	0.9	410	10	BT039948	BT039948 MR4-NPT014
588	17	0.9	360	9	AV194655	AV194655	661	17	0.9	411	10	BF188721	BF188721
589	17	0.9	360	10	C50400	C50400 yu1j1	662	17	0.9	412	9	AM662635	AM662635 h133h04.x
590	17	0.9	360	10	D69839	D69839 CELK02G4F	663	17	0.9	414	9	AM519964	AM519964
591	17	0.9	360	10	D73983	D73983 CELK0672F	664	17	0.9	416	9	AA461730	AA461730
592	17	0.9	360	10	D74063	D74063 CELK04E1F	665	17	0.9	418	10	BM286810	BM286810
593	17	0.9	360	10	D74690	D74690 CELK04B5F	666	17	0.9	418	12	AQ206779	AQ206779
594	17	0.9	360	12	AQ283182	AQ283182 RPTC111-80	667	17	0.9	419	10	BF052856	BF052856
595	17	0.9	365	9	AM928766	AM928766 EST337554	668	17	0.9	419	10	BF707526	BF707526 248770.MA
596	17	0.9	365	10	BM217713	BM217713 C0901B02-	669	17	0.9	420	9	A1440690	A1440690 sa67g12.y
597	17	0.9	368	10	BM272236	BM272236 i942a09.x	670	17	0.9	421	10	BF071522	BF071522 st62b10.y
598	17	0.9	368	10	R04076	R04076 pk19a12.s1	671	17	0.9	422	9	AA830935	AA830935
599	17	0.9	369	10	BG273315	BG273315 OV2.25.H0	672	17	0.9	422	9	AM754948	AM754948 PC08G02.P
600	17	0.9	369	10	BE520504	BE520504 M13A10XP	673	17	0.9	422	10	BM390569	BM390569
601	17	0.9	369	10	BF501010	BF501010 AT16194.5	674	17	0.9	422	10	BE673931	BE673931

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	602460969F1 NIH_MGC_20	Homo sapiens cDNA clone IMAGE:457753 5'	602460969F1	NIH_MGC_20	IMAGE:457753 5'	human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 928)	NIH-MGC http://mgc.nci.nih.gov/	unpublished (1999)	Contact: Robert Strausberg, Ph.D. Email: c9abs-remail.nih.gov Tissue Procurement: ATCC/CDT/DP CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Sequencing by: Inceye Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov plate: L10M1292 row: 1 column: 18 High quality sequence stop: 747.

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	602460969F1 NIH_MGC_20	Homo sapiens cDNA clone IMAGE:457753 5'	602460969F1	NIH_MGC_20	IMAGE:457753 5'	human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	1 (bases 1 to 928)	NIH-MGC http://mgc.nci.nih.gov/	unpublished (1999)	Contact: Robert Strausberg, Ph.D. Email: c9abs-remail.nih.gov Tissue Procurement: ATCC/CDT/DP CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Sequencing by: Inceye Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov plate: L10M1292 row: 1 column: 18 High quality sequence stop: 747.

FEATURES

source
1. .928
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:457753"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT

234 a 245 c 302 g 147 t

Query Match

Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

570 CTGCTGAACAGTATGAGAACGA 592
661 CTGCTGAACAGTATGAGAACGA 683

RESULT

2
LOCUS A0701777/c 553 bp DNA linear GSS 07-JUL-1999
DEFINITION HS_2134_A2.C12_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2134 Col-24 Row-E, DNA sequence.

ACCESSION

A0701777
A0701777.1 GI:5411203

KEYWORDS

GSS.

SOURCE

human.

ORGANISM

human.

REFERENCE

1 (bases 1 to 553)

AUTHORS

Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL

Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE

99380589

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 2134 row: E column: 24
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 553.

FEATURES

source
1. .553
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-2134 Col-24 Row-E"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"

BASE COUNT

146 a 97 c 105 g 197 t 8 others

ORIGIN

Query Match 1.1%; Score 22; DB 12; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 737 TACACAATGTAGACCAAGAG 758

```

Db      531 TACACAAATGTAGACCAAG 510
|||||
RESULT  3
LOCUS   CENS0363A
DEFINITION Tetradon nigroviridis genome survey sequence PUC-ori end of clone
          215173 of library G from Tetradon nigroviridis, genomic survey
          sequence.
ACCESSION AL228519
VERSION   AL229519.1 GI:7888514
KEYWORDS  GSS; genome survey sequence.
SOURCE    Tetradon nigroviridis.
ORGANISM  Tetradon nigroviridis.
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
          Acanthomorpha; Acanthopterygii; Percomorphia; Tetraodontiformes;
          Tetraodontidae; Tetradon.
REFERENCE 1 (bases 1 to 909)
AUTHORS   Roest-Crollius,H., Jallion,O., Dasilva,C., Fzames,C., Fisher,C.,
          Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
          Weissenbach,J.
          Characterization and repeat analysis of the compact genome of the
          freshwater pufferfish Tetradon nigroviridis
          Unpublished
          2 (bases 1 to 909)
          Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
          Bernot,A., Fzames,C., Wincker,P., Brothier,P., Quetier,F.,
          Saurin,W. and Weissenbach,J.
          Human gene number estimate provided by genome wide analysis using
          Tetradon nigroviridis DNA sequence
          Unpublished
          3 (bases 1 to 909)
          Genoscope.
          Direct Submission
          Submitted (12-APR-2000) to the EMBL/Genbank/DBJ databases
          This sequence is a single read and was generated as part of a large
          scale clone-end sequencing project of the Tetradon nigroviridis
          genome. For more information, please take a look at
          http://www.genoscope.cns.fr/Tetradon.
FEATURES
    source
        1..909
        /organism="Tetradon nigroviridis"
        /db_xref="taxon:99883"
        /clone="215123"
        /clone_1lb="G"
        /note="Genoscope sequence ID : C0AG215AE12SP1-end :
        PUC-ori"
BASE COUNT 226 a 205 c 243 g 228 t 7 others
ORIGIN
Query Match 1.1%; Score 22; DB 12; Length 909;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1257 CAGACAAAAAATATGAGCAAG 1278
|||||
Db      567 CAGACAAAAAATATGAGCAAG 588

```

```

RESULT  4
LOCUS   AL632585
DEFINITION AL632585 XGC.gastrula silurana tropicalis cDNA Tgas021e05 5',
          mRNA sequence.
ACCESSION AL632585
VERSION   AL632585.1 GI:16602068
KEYWORDS  EST.
SOURCE    western clawed frog.
          Silurana tropicalis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
          Xenopodinae; Silurana.
          1 (bases 1 to 389)
REFERENCE 1
AUTHORS   Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
          Sanger Xenopus tropicalis EST project 2001 (10_2001)
          Unpublished (2001)
          Contact: Huckle E
          Sanger Centre
          Hinxton, Cambridgeshire, CB10 1SA, UK
          Email: tropesanger.ac.uk
          Sanger Xenopus tropicalis EST project 2001
          TROPICALIS_SEQUENCE_ID: Tgas021e05.sp6
          Sequencing primer: SP6
          This sequence is from a Xenopus Gene Collection (XGC) library
          constructed by Aaron M. Zorn.
FEATURES
    source
        1..389
        /organism="Silurana tropicalis"
        /db_xref="taxon:8364"
        /clone="Tgas021e05"
        /clone_1lb="XGC-gastrula"
        /dev_stage="gastrula (stages 10.5-13 mixed)"
        /lab_host="Escherichia coli XL1-blue"
        /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
        was oligo dt primed from 5ug of poly A+ RNA from stages
        10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
        into pCS107 with EcoRI at the 5' end and NotI at the 3'
        end."
BASE COUNT 93 a 110 c 104 g 82 t
ORIGIN
Query Match 1.1%; Score 21; DB 9; Length 389;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1429 TATGATGACCGACATGATGAA 1449
|||||
Db      312 TATGATGACCGACATGATGAA 332

```

```

RESULT  5
LOCUS   AI397462/c
DEFINITION AI397462 422 bp mRNA linear EST 07-JUN-2001
          FB18B08.x1 zebrafish Washu MP1MG EST Danio rerio cDNA clone
          IMAGE:3712215.3, similar to gb:57527 COLLAGEN ALPHA 1(VIII) CHAIN
          PRECURSOR (HUMAN); mRNA sequence.
ACCESSION AI397462
VERSION   AI397462.1 GI:4227355
KEYWORDS  EST.
SOURCE    zebrafish.
          Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
          1 (bases 1 to 422)
REFERENCE 1
AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
          ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
          ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
          Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
          Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
          and Wilson,R.
          Washu zebrafish EST Project 1998
          Unpublished (1998)
          Contact: Stephen L. Johnson
          Washington University School of Medicine
          4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
          Tel: 314 286 1800
          Fax: 314 286 1810
          Email: zbrafish@wustl.edu
          cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
          Matthew Clark. DNA Sequencing by: Washington University Genome
          Sequencing Center Clone distribution: Genome Systems, St. Louis,

```

Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and Ressourcenzentrum/Primatendbank, Berlin, Germany (web address: www.tzpd.de)
Seq primer: T7 ET from Amersham
High quality sequence stop: 415
POLYA-No.

FEATURES

SOURCE

Location/Qualifiers
1..422
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="IMAGE:3712215"
/clone_11p="Zebrafish Washu MPIMG ESF"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield stage embryos"
/lab_host="XLI-blue MRF"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st strand cDNA was primed with a Not I - oligo(dT)15 primer [5,pgACTACTTCTAGATCGGAGCGGCCCGCTTTTCTTTTCTT3']; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab, ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

BASE COUNT 120 a 88 c 97 g 117 t
ORIGIN

Query Match 1.1%; Score 21; DB 9; Length 422;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1757 GGCTCTACAGAAATGGAGA 1777
|||||
Db 236 GGCTCTACAGAAATGGAGA 216

RESULT 6
AL629478 704 bp mRNA linear EST 02-NOV-2001
LOCUS AL629478
DEFINITION AL629478 XGC-gastrula silurana tropicalis cDNA clone Tgas018104 5', mRNA sequence.
ACCESSION AL629478
VERSION AL629478.1 GI:16598961
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Silurana.
1 (bases 1 to 704)
Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
Unpublished (2001)
Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: tropesanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: Tgas018104.sp6
Sequencing primer: SP6

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.
Location/Qualifiers
1..704
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="Tgas018104"
/clone_11p="XGC-gastrula"
/dev_stage="gastrula (stages 10-5-13 mixed)"
/lab_host="Escherichia coli XLI-blue"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

BASE COUNT 153 a 213 c 188 g 150 t
ORIGIN

Query Match 1.1%; Score 21; DB 9; Length 704;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 TATGATGACGACATGATGAA 1449
|||||
Db 318 TATGATGACGACATGATGAA 338

RESULT 7
CNS06N80 1025 bp DNA linear GSS 17-JUN-2001
LOCUS T3 end of clone AU00A010D02 of library AU00A from strain CBS 3082
DEFINITION of Saccharomyces kluyveri, genomic survey sequence.
ACCESSION AL406462
VERSION AL406462.1 GI:12170499
KEYWORDS GSS.
SOURCE Saccharomyces kluyveri.
ORGANISM Saccharomyces kluyveri
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS

1 (bases 1 to 1025)
Neuveglise,C., Bon,E., Lepingle,A., Wincker,P., Artiguenave,F., Gaillardin,C. and Casaregola,S.
Genomic exploration of the hemiascomycetous yeasts: 9. Saccharomyces kluyveri
FEBS Lett. 487 (1), 56-60 (2000)
20584719

2 (bases 1 to 1025)
Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potter,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
20584711

3 (bases 1 to 1025)
Genoscope.
Direct Submission
Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: segref@genoscope.cns.fr Web: www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of

COMMENT

the other extremity of this insert.

FEATURES
source
1..1025
/organism="Saccharomyces kluyveri"
/strain="CBS 3082"
/db_xref="taxon:4934"
/clone="AU00A010D02"
/clone_11b="AU00AA"
/note="end : T3"
<30..>1025

misc-feature

/note="similar to Saccharomyces cerevisiae ORF YGR233c [PH081 : cyclin-dependent kinase inhibitor]"
/evidence-not_experimental

BASE COUNT 348 a 191 c 205 g 280 t 1 others
ORIGIN

Query Match 1.1%; Score 21; DB 12; Length 1025;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1460 ACTTATGGAAGTGCAGCAAT 1480
DB 590 ACTTATGGAAGTGCAGCAAT 610

RESULT 8

CNS06NM7/c 1078 bp DNA linear GSS 04-JUL-2001
DEFINITION T7 end of clone AU00A015C12 of library AU00A from strain CBS 3082
of Saccharomyces kluyveri, genomic survey sequence.

ACCESSION AL406949
VERSION AL406949.1 GI:12171649
KEYWORDS GSS.

SOURCE

ORGANISM Saccharomyces kluyveri.
Saccharomyces kluyveri.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

REFERENCE 1 (bases 1 to 1078)
Neueglise, C., Bon, E., Lepingle, A., Winker, P., Attienave, F.,
Gallardin, C., and Casaregola, S.
Genomic exploration of the hemiascomycetous yeasts: 9.
Saccharomyces kluyveri

TITLE

JOURNAL FEMS Lett. 487 (1), 56-60 (2000)
MEDLINE 20584719

REFERENCE

AUTHORS 2 (bases 1 to 1078)
Souciet, J. L., Aigle, M., Attienave, F., Blandin, G.,
Boitlin-Fukuhara, M., Bon, E., Brothier, P., Casaregola, S.,
de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Lorente, B.,
Malpertuy, A., Neueglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,
Winker, P., and Weissenbach, J.

TITLE

JOURNAL FEMS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711

3 (bases 1 to 1078)
Genoscope.

TITLE

JOURNAL Direct Submission
MEDLINE Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

REFERENCE

AUTHORS This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
anastus, Debaryomyces hansenii var. hansenii, Pichia sorbicola,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

FEATURES

Location/Qualifiers

source

1..1078
/organism="Saccharomyces kluyveri"
/strain="CBS 3082"
/db_xref="taxon:4934"
/clone="AU00A015C12"
/clone_11b="AU00AA"
/note="end : T7"
complement(<1..>1074)

misc-feature

/note="similar to Saccharomyces cerevisiae ORF YGR233c [PH081 : cyclin-dependent kinase inhibitor]"
/evidence-not_experimental

BASE COUNT 300 a 218 c 214 g 345 t 1 others
ORIGIN

Query Match 1.1%; Score 21; DB 12; Length 1078;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1460 ACTTATGGAAGTGCAGCAAT 1480
DB 739 ACTTATGGAAGTGCAGCAAT 719

RESULT 9

BF174868 237 bp mRNA linear EST 23-MAR-2001
LOCUS MYE3951 Myeloma (MYE) cDNA library Homo sapiens cDNA, mRNA
DEFINITION sequence.

ACCESSION BF174868
VERSION BF174868.1 GI:13441082
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 237)
Claudio, J. O., Tang, H., Khan, E. M., Vorralla, M., Li, Z., Cukerman, E.,
Francisco-Pabalan, O., Ilew, C. C. and Stewart, A. K.

The transcriptional phenotype of myeloma cells
Unpublished (2000)
Contact: A. Keith Stewart, M.D.
Oncology Research
University Health Network
610 University Ave., 5-126, Toronto, Ontario, M5G 2M9, Canada
Tel: (416) 946-4639
Fax: (416) 946-6546
Email: k.stewart@utoronto.ca

TITLE

JOURNAL PCR Primers
COMMENT FORWARD: 5'-GCCAGCTCGCAATTAACCTCCTAAGG-3'
BACKWARD: 5'-CCAGTGAATTTGTAATACGACTCCTAAGGCG-3'
Seq primer: 5'-GAATTTAACCTCCTAAGG-3'.

FEATURES

source

Location/Qualifiers

1..237
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_11b="Myeloma (MYE) cDNA library"
/sex="male"
/tissue_type="Blood"
/cell_type="myeloma"
/dev_stage="multiple myeloma"

/note="Vector: lambda zap Express; Site1: EcoRI; Site2:
XhoI; Myeloma cells from multiple myeloma patients; bone
marrow were purified by magnetic cell sorting. mRNA were
restriction and an oligo d(T)18 primer containing XhoI
restriction site was used to prime first strand synthesis
using M-MV reverse transcriptase. To protect the cDNAs
from XhoI digestion in subsequent cloning step, the
nucleotide analogue 5-methyl-dCTP was added to the
nucleotide mixture and [a-32P]dATP was added to monitor
the quantity and quality of first strand synthesis. After
second-strand synthesis and blunting of cDNA termini,
EcoRI adapters were ligated, followed by kinase treatment

and digestion with XhoI. The cDNAs were then size-fractionated using Sephacryl S-500 column and then ligated into EcoRI and XhoI digested Lambda Zap Express vector. The ligation product was packaged using Gigapack II packaging extract. The library had primary titre of approx. 1x10⁶. Clones from the primary library were randomly selected for single pass sequencing.

BASE COUNT 103 a 35 c 54 g 45 t

Query Match 1.0%; Score 20; DB 10; Length 237;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAAGAAA 338
|||||
Db 191 AAAAGCAACCAAGCAAGAAA 210

RESULT 10
AA317576 287 bp mRNA linear EST 19-APR-1997
LOCUS EST19524 Retina II Homo sapiens cDNA 5' end similar to DNA binding
DEFINITION protein PO-GA, mRNA sequence.
ACCESSION AA317576
VERSION AA317576.1 GI:1969905
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 287)
Adams,M.D., Kesteven,A.R., Fleischmann,R.D., Fuldner,R.A., Bolt,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wei,C., Clayton,R.A., Cline,T.R., Colton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodex,A., Guelh,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Uteback,T.R., Weidman,J.F., Li,Y., Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weisner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Pannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
COMMENT Other-ESTs: TH0171722

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018689056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/cdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 287

FEATURES
source /organism="Homo sapiens"
/db_xref="ATCC (inhost):118036"
/db_xref="taxon:9606"
/clone_lib="Retina II"
/sex="mixed"

/dev_stage="adult"
/note="Organ: retina; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 113 a 57 c 64 g 52 t 1 others

Query Match 1.0%; Score 20; DB 9; Length 287;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAAGAAA 338
|||||
Db 243 AAAAGCAACCAAGCAAGAAA 262

RESULT 11
BB551207 295 bp mRNA linear EST 01-AUG-2000
LOCUS BB551207 RIKEN full-length enriched, 2 days pregnant adult female
DEFINITION oviduct Mus musculus cDNA clone E230026H21 3' similar to D10712 Mus
musculus mRNA for nedd-1 protein, mRNA sequence.
ACCESSION BB551207
VERSION BB551207.1 GI:9637573
KEYWORDS EST.
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 295)
Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya,T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamaoka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Kono,H., et al.)
Unpublished (2000)
Contact: Yoshihide Hayashizaki.
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsunumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermocyclization of labile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Katsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES
source /organism="Mus musculus"
/db_xref="taxon:10090"
/clone="E230026H21"

Matches	20: Conservative	0: Mismatches	0: Indels	0: Gaps
QY	325 AACCAAGCAGAAATCTGA	344		
Db	242 AACCAAGCAGAAATCTGA	261		

Accession	Definition	LOCUS	Result 14
BM173589	900943 Avicennia marina leaf cDNA library Avicennia marina cDNA clone AM00943.5, similar to NADH dehydrogenase subunit 5 (AF085145) of Schistosoma mansoni, mRNA sequence.	402 bp	EST 04-DEC-2001
BM173589			

ORGANISM	Avicennia marina.
SOURCE	Avicennia marina.
KEYWORDS	EST.
VERSION	BM173589.1
ACCESSION	GI:17313152
BM173589	

REFERENCE	1 (bases 1 to 402)
AUTHORS	Parani,M., Mehta,P., Sivaprakash,K.R. and Parida,A.
TITLE	Expressed sequence tags from the mangrove species Avicennia marina
JOURNAL	Unpublished (2000)
COMMENT	Contact: Parani M / Parida A

II Cross Street, Taramani Institutional Area, Chennai 600 113,
India
Tel: 91-44-2351319
Fax: 91-44-2351319
Email: mangrovegenes@msrf.res.in
BLAST search in BLASTx (Non-redundant) using default parameters as
on November 6th 2001
Seq primer: M13 Reverse: -3'-TGTACGACCTGTTAGTTTGG-5'

FEATURES	Location/Qualifiers
source	1. .402

```

/organism="Avicennia marina"
/strain="Pichavaram"
/db_xref="taxon:82927"
/clone="Am900943"
/clone_lib="Avicennia marina leaf cDNA Library"
/sex="Hermaphrodite"
/tissue_type="leaf"
/dev_stage="1 year old plant"
/note="A one-year-old plant from the natural mangrove
habitat in Pichavaram, Tamil Nadu, India was collected
and brought to the M.S.Swaminathan Research Foundation.
The whole plant was treated with the nutrient solution
supplemented with 500 mM NaCl for 48 hours. Poly(A)+mRNA
from the leaf tissue was purified and cDNA was prepared
using Superscript Lambda System (Life Technologies, Cat.
No. 19643-014). The cDNAs were size fractionated over
SizeSep-400 spun column (Amersham-Pharmacia Biotech Cat
No. 27-5105-01) and cloned in 5'Sal I - 3' Not I of
pSPORT1 (Life Technologies Cat. No. 15383-011). The ESTs
were sequenced from the 5' end using M13/pUC18 reverse
primer in an Automated Sequencer (ABI310). Applied
Biosystems) and submitted after editing to remove the
vector and adapter sequences. BLAST search in BLASTX
(Non-redundant) was carried out on 7th August, 2000 using
default parameters and the results reported under
putative identification/comment"

```

Query Match	1.08;	Score 20;	DB 10;	Length 402;
-------------	-------	-----------	--------	-------------

Best Local Similarity	100.0%	Pred. No. 39;			
Matches	20; Conservative	0; Mismatches	0; Indels	0; Gaps	0;

QY 1281 GTGAAAGATGGCTGCTGCT 1300

Db 341 GTGAAAGATGGCTGCTGCT 360

RESULT 15	LOCUS	DEFINITION
BF016982	405 bp	mRNA
BF016982	405 bp	linear
IMAGE:3471952	5'	similar to SW:ND1_MOUSE P33215 NEDD1 PROTEIN ; ,
mRNA sequence.		

ACCESSION	BF016982
VERSION	BF016982.1
KEYWORDS	GI:10748314
SOURCE	EST.
ORGANISM	house mouse.
	Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Schiornath; Muridae; Mus
1 (Pages 1 to 405)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CCAP),
Tumor Gene Index
Unpublished (1997)
Other_ESTRs: ux19c09.x1

Contact: Robert Strassberg, Ph.D.
Email: cgabs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:339332
Seq primer: -40RP from Gibco
High quality sequence stop: 344.

FEATURES	Location/Qualifiers
source	1. .405

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:341952"
/clone_lib="Scares_thymus_2NBMT"
/sex="male"

```

```

/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/node="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dn) primer [5',
TGTTCACCAATCTGAGTGGGAGCGCGCGTGTATTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT733 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."

```

BASE COUNT	124 a	101 c	88 g	92 t
ORIGIN				

Query Match	1.0%;	Score 20;	DB 10;	Length 405;
Best Local Similarity	100.0%;	Pred. No. 39;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	325	AACCAAGCAAGAAATCTTGA	344
Db	283	AACCAAGCAAGAAATCTTGA	302

RESULT 16
AQ445724

LOCUS	AO445724	406 bp	DNA	linear	GSS 16-DEC-1999
DEFINITION	nbx0054j07f	CUGI	Rice BAC Library	Oryza sativa	genomic clone
	nbx0054j07f,	DNA sequence.			

ACCESSION	AQ445724
VERSION	AQ445724.2
KEYWORDS	GSS.
SOURCE	Oryza sativa.

Oryza sativa.

ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrharioideae; Oryzaceae; Oryza.
1 (bases 1 to 406)

REFERENCE
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
On Dec 15, 1999 this sequence version replaced gi:4574947.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4298
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TATATCGACTCCTATAGCG
Class: BAC ends
High quality sequence stop: 376.

FEATURES
source
1..406
Location/Qualifiers
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipondare"
/db_xref="taxon:4530"
/clone="nxb0054j07f"
/clone.lib="CGI Rice BAC library"
/tissue.type="leaf"
/lab_host="E. coli DH10B"
/note="Vector: pBel0AC11; Site_1: HindIII; Site_2:
HindIII; Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipondare variety. The
library contains 36,864 clones with an average insert size
of 128.5 kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation of a particular sequence
with a probability of 99.9%. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."

BASE COUNT
129 a 77 c 64 g 136 t

ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 406;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 636 CAAGTAATACCTGCTTAC 655
|||||
Db 219 CAAGTAATACCTGCTTAC 238

RESULT 17
BG017796 440 bp mRNA linear EST 30-JAN-2001
LOCUS BG017796
DEFINITION db10f11.y1 NICHD XGC Spl Xenopus laevis cDNA clone IMAGE:4174220
5, similar to TR:09016 Q90Y16 MDJ7.;, mRNA sequence.
ACCESSION BG017796 GI:12472380
VERSION BG017796.1
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

REFERENCE
1 (bases 1 to 440)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS
TITLE
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapsb-r@mail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LNL at: info@image.llnl.gov
Seq primer: -40RP from Gibco.

FEATURES
source
1..440
Location/Qualifiers
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IMAGE:4174220"
/clone.lib="NICHD XGC Spl"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: spleen; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.4 kb. Constructed by Life
Technologies."

BASE COUNT
136 a 83 c 99 g 122 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 440;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1306 TGGCTTTGCAATGTGGAA 1325
|||||
Db 335 TGGCTTTGCAATGTGGAA 354

RESULT 18
BB824552 442 bp mRNA linear EST 19-NOV-2001
LOCUS BB824552
DEFINITION BB824552 RIKEN full-length enriched, mammary gland RCB-0526
Jyg-MC(A) CDNA Mus musculus cDNA clone G830033120 3', mRNA
sequence.
ACCESSION BB824552
VERSION BB824552.1 GI:17002795
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 442)

REFERENCE
Akimura, T., Arai, K., Carninci, P., Furuno, M., Hanagaki, T.,
Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K., Ishii,
Y., Ito, M., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Matsuyama, T.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,
Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K.,
Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa,
A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T.,
Watanuki, A., Yasunishi, A., Yamatsuta, M. and Hayashizaki, Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.
2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.go.jp,

URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 Waga, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
 e mouse tissues.

FEATURES

source location/Qualifiers
 1..442
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="G830033i20"
 /clone_lib="RIKEN full-length enriched, mammary gland
 RCB-0526 Jyg-MC(A) cDNA"
 /tissue_type="mammary gland"
 /cell_line="RCB-0526 Jyg-MC(A)"
 BASE COUNT 138 a 93 c 109 g 102 t
 ORIGIN

Query Match 1.0%; Score 20; DB 9; Length 442;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 325 ACCACAGCAAGAAATCTTGA 344
 ||||||||||||||||
 Db 408 AACCAAGCAAGAAATCTTGA 427

RESULT 19
 BE566388 443 bp mRNA linear EST 15-AUG-2000
 LOCUS 601339984F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3682221 5',
 DEFINITION mRNA sequence.
 ACCESSION BE566388
 VERSION BE566388
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 443)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
 Plate: L1CM365 row: g column: 22
 High quality sequence stop: 441.
 Location/Qualifiers
 1..443
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

FEATURES

source
 1..443
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/clone="IMAGE:3682221"
 /clone_lib="NIH_MGC_53"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: bladder; Vector: pDNR-LIB (Clontech);
 Site_1: SfiI (ggcgctcgcc); Site_2: SfiI (ggccatagcc
); Double-stranded cDNA was prepared from cell line RNA.
 5' and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CAGCGCATATGCGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGCGCGCGCGCATG-dT(30)BN-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA)."
 BASE COUNT 167 a 91 c 100 g 84 t 1 others
 ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 443;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 319 AAAGCAAGCAAGCAAGAA 338
 ||||||||||||||||
 Db 274 AAAGCAAGCAAGCAAGAA 293

RESULT 20
 BG942702 446 bp mRNA linear EST 11-JUN-2001
 LOCUS ax28e04.x1 Proliferating Human Erythroid Cells (LCB:ax library)
 DEFINITION Homo sapiens cDNA clone ax28e04 random, mRNA sequence.
 ACCESSION BG942702
 VERSION BG942702.1 GI:14342074
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 446)
 Gubin, A.N., Nforge, J.M., Bouffard, G.G. and Miller, J.L.
 Gene expression in proliferating human erythroid cells
 Genomics 59 (2), 168-177 (1999)
 99339981
 Contact: Jeffery L. Miller
 Laboratory of Chemical Biology
 National Institute of Diabetes and Digestive and Kidney Diseases
 Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
 20892, USA
 Tel: 301 402 2373
 Fax: 301 435 5148
 Email: jm7fe@nih.gov
 DNA Sequencing and analyses by National Institutes of Health
 Intramural Sequencing Center (NISC).
 Plate: 28 row: e column: 04
 Seq primer: -21M13 forward primer (ABI).
 Location/Qualifiers
 1..446
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ax28e04"
 /clone_lib="Proliferating Human Erythroid Cells (LCB:ax
 library)"
 /sex="unknown"
 /tissue_type="blood"
 /cell_line="Erythroid Cells"
 /cell_line="Primary Culture of Peripheral Blood
 Mononuclear Cells"
 /dev_stage="Progenitor; EPO responsive CD11++++"
 /lab_host="SOLR"
 /note="Organ: Blood; Vector: Lambda ZAP II; Site_1: EcoRI;
 Site_2: EcoRI; 65,000 proliferating erythroid cells from

the buffy coat of a blood donation were obtained by flow cytometric separation after a 5-day culture period in the presence of erythropoietin. Total RNA was purified from the sorted cell population using TRIzol reagent. RNA (0.3 ug) was converted into double stranded cDNA using Clontech's Capfinder cDNA Library Construction Kit (Clontech) according to the manufacturer's protocol and cloned into EcoRI digested lambda zap II vector (Stratagene). The phage library was amplified once prior to in vivo excision in SOLR cells. Individual colonies were grown, and the cDNA inserts were sequenced in high throughput (NIH Intramural sequencing center <http://www.nisc.nih.gov/>).

BASE COUNT 171 a 85 c 102 g 88 t
ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 446;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAGCACCAGCAGAGAA 338
Db 260 AAAGCACCAGCAGAGAA 279

RESULT 21

LOCUS BB750480 450 bp mRNA linear EST 16-OCT-2001
DEFINITION BB750480 RIKEN full-length enriched, pooled tissues, cerebellum,
etc. Mus musculus cDNA clone G130003M12.3', mRNA sequence.

ACCESSION BB750480
VERSION BB750480.1 GI:16154716

KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 450)

Akimura,T., Arikawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,
Y., Ito,M., Kawai,T., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shingawa,A., Shireki,T., Sogabe,Y., Suzuki,H., Tagawa,
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Wataniki,A., Yasunishi,A., Muramatsu,M., and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suicho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

EMAIL: genome-resgsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>,
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,
M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

Waki,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,
Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.
e mouse tissues.

FEATURES
source
Location/Qualifiers

1. 450
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G130003M12"
/clone.lib="RIKEN full-length enriched, pooled tissues,
cerebellum, etc."
/note="pooled tissues : (tissue_type=cerebellum,
dev_stage=16 days neonate, sex=mixed),
(tissue_type=cerebellum, dev_stage=0 day neonate,
sex=mixed), (tissue_type=hippocampus, dev_stage=adult,
sex=adult), (tissue_type=whole body, dev_stage=9 days
embryo, sex=mixed), (tissue_type=lung, dev_stage=13 days
embryo, sex=mixed)"
BASE COUNT 141 a 94 c 106 g 109 t
ORIGIN

Query Match 1.0%; Score 20; DB 9; Length 450;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 325 AACCAAGCAAGAAATCTTGA 344
Db 416 AACCAAGCAAGAAATCTTGA 435

RESULT 22

LOCUS AA280963 493 bp mRNA linear EST 15-AUG-1997
DEFINITION z597d12.r1 NCI-CGAP GCBI Homo sapiens cDNA clone IMAGE:711575.5'
similar to gb:U23320 ACTIVATOR 1 140 KD SUBUNIT (HUMAN);, mRNA
sequence.

ACCESSION AA280963
VERSION AA280963.1 GI:1923644
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 493)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lml.gov) for further information.
Seq primer: 28m13 rev2 ET from Amersham
High quality sequence stop: 469.

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lml.gov) for further information.
Seq primer: 28m13 rev2 ET from Amersham
High quality sequence stop: 469.

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lml.gov) for further information.
Seq primer: 28m13 rev2 ET from Amersham
High quality sequence stop: 469.

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.lml.gov) for further information.
Seq primer: 28m13 rev2 ET from Amersham
High quality sequence stop: 469.

[5'-TGTACCAATCTGAGTGGAGCGCCGCTCATTTTCTTTTCTTT-3'
]. Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pRT3 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 188 a 94 c 115 g 96 t
 ORIGIN

Query Match 1.0%; Score 20; DB 9; Length 493;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAAGAA 338
 Db 274 AAAAGCAACCAAGCAAGAA 293

RESULT 23 526 bp mRNA linear EST 04-MAY-2000
 LOCUS AW762116
 DEFINITION ur53h07.y1 NCI-CGAP_Mam3 Mus musculus cDNA clone IMAGE:3154045 5'
 (similar to gb:D10712 Mouse mRNA for nedd-1 protein, complete cds
 (MOUSE)), mRNA sequence.

ACCESSION AW762116 GI:7694034
 VERSION AW762116
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 526)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other-ESTs: ur53h07.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrived by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 image.llnl.gov/image/html/iresources.shtml

JOURNAL
 AUTHORS
 TITLE
 COMMENT

FEATURES
 source
 MGI:1056801
 Seq primer: -40RP from Gibco
 High quality sequence stop: 405.
 Location/Qualifiers

1..526
 /organism="Mus musculus"
 /strain="129 - C57/B6 - FVBN"
 /db_xref="taxon:10090"
 /clone_image="3154045"
 /clone_lib="NCI-CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: PCMV-SPORT6, site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."

BASE COUNT 159 a 130 c 123 g 114 t
 ORIGIN

Query Match 1.0%; Score 20; DB 9; Length 526;
 Best Local Similarity 100.0%; Pred. No. 41;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 325 AACCAAGCAAGAAATCTGA 344
 Db 299 AACCAAGCAAGAAATCTGA 318

RESULT 24 526 bp DNA linear GSS 13-JUL-1999
 LOCUS AO710693
 DEFINITION HS_5344_A2_E10_T7A RCI-11 Human Male BAC Library Homo sapiens
 genomic clone Plate=920 Col=20 Row=I, DNA sequence.
 ACCESSION AO710693
 VERSION AO710693.1 GI:5460009
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 526)
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
 99380589
 CONTACT: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieterdejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
 or from Research Genetics (info@resgen.com). BAC end web server:
 http://www.htsc.washington.edu
 Plate: 920 row: I column: 20
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 526.
 Location/Qualifiers

1..526
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_plate="920 Col=20 Row=I"
 /clone_lib="RCI-11 Human Male BAC Library"
 /sex="male"
 /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
 male blood DNA was isolated from one randomly chosen donor
 and partially digested with a combination of EcoRI and
 EcoRI-MethyIase. Size selected DNA was cloned into the
 pBACe3.6 vector at EcoRI sites"
 BASE COUNT 123 a 103 c 113 g 185 t 2 others
 ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 526;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1699 CTCATGATGACCTACTGT 1718
 Db 431 CTCATGATGACCTACTGT 450

RESULT 25 535 bp DNA linear GSS 20-FEB-2001
 LOCUS AZ824983
 DEFINITION 2M0099J24R Mouse 10kb plasmid UUCIM library Mus musculus genomic

clone UUGC2M009J24 R, DNA sequence.
 ACCESSION A2824983
 VERSION A2824983.1 GI:12994891
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 535)
 REFERENCE Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamll, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausen, A. and Wright, D., Weiss, R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0099 row: J column: 24
 Seq primer: CACACGAGAAACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 535.
 Location/Qualifiers
 1..535
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M009J24"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (911473211419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 164 a 130 c 120 g 121 t
 ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 535;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 AACCAAGCAAGAAATCTTGA 344
 ||||||||||||||||||
 Db 301 AACCAAGCAAGAAATCTTGA 320

RESULT 26
 BG723460 563 bp mRNA linear EST 08-MAY-2001
 LOCUS

602694145F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4826218 5',
 mRNA sequence.
 ACCESSION BG723460
 VERSION BG723460.1 GI:14002647
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 563)
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki Toshitsuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1AM10740 row: J column: 11
 High quality sequence stop: 406.
 Location/Qualifiers
 1..563
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4826218"
 /clone_lib="NIH_MGC_97"
 /lab_host="DH10B"
 /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.2 kb and normalized to 80F 5. This is a primary library enriched for full-length clones and constructed using the cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC library."

BASE COUNT 201 a 124 c 129 g 109 t
 ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 563;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAGAGAA 338
 ||||||||||||||||||
 Db 287 AAAAGCAACCAAGCAGAGAA 306

RESULT 27
 BG503218 587 bp mRNA linear EST 27-MAR-2001
 LOCUS 60251318F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4658644 5',
 mRNA sequence.
 ACCESSION BG503218
 VERSION BG503218.1 GI:13464735
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 587)
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LNCM1453 row: d column: 05
 High quality sequence stop: 587.

FEATURES

source

```
1.587
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4658644"
/clone_1lb="NIH_MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
Still (ggcgccctggcc); Site_2: Still (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCTAGAGCGGCGGCCGACATG-dt(30)-BN-3'
(sequence: 5'-ATTCTAGAGCGGCGGCCGACATG-dt(30)-BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
```

BASE COUNT 237 a 112 c 126 g 112 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 587;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAGCAACCAAGCAGAGAA 338
 ||||||||||||||||
 Db 268 AAAGCAACCAAGCAGAGAA 287

RESULT 28 588 bp mRNA linear EST 03-APR-2001
 BG539230 602567826F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4692414 5',
 LOCUS mRNA sequence.
 DEFINITION
 ACCESSION BG539230
 VERSION BG539230.1 GI:13531463
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 588)
 AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LNCM1513 row: c column: 07
 High quality sequence stop: 586.

FEATURES

source

```
1.588
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4692414"
```

```
/clone_1lb="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
Still (ggcgccctggcc); Site_2: Still (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGCGGCCGACATG-dt(30)-BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
```

BASE COUNT 227 a 113 c 130 g 118 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 588;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAGCAACCAAGCAGAGAA 338
 ||||||||||||||||
 Db 289 AAAGCAACCAAGCAGAGAA 308

RESULT 29 642 bp mRNA linear EST 30-SEP-1997
 AA606845
 LOCUS vM94h07.r1 Knowles Solter mouse blastocyst B1 mus musculus cDNA
 DEFINITION clone IMAGE:1005949 5' similar to gb:U0712 Mouse mRNA for nedd-1
 protein, complete cds (MOUSE);, mRNA sequence.
 ACCESSION AA606845
 VERSION AA606845.1 GI:2455738
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 642)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Stepcoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:570165

High quality sequence stop: 473.
 Location/Qualifiers

FEATURES

source

```
1.642
/organism="Mus musculus"
/strain="B6D2 F1/J"
/db_xref="taxon:10090"
/clone="IMAGE:1005949"
/clone_1lb="Knowles Solter mouse blastocyst B1"
/tissue_type="blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"
/note="Organ: embryo; Vector: pSPORT; Site_1: NotI;  

Site_2: SalI; Cloned unidirectionally from mRNA prepared  

from 800 blastocysts. Primer: SalI(dT):  

5'-CGGTCACCGTCGACCGCTTTTCTTTT-3'. cDNAs were  

cloned into the NotI/SalI sites of a pSPORT vector (Life  

Technologies). Two different size selections: B1 (larger
```


BASE COUNT 196 a 138 c 143 g 165 t
ORIGIN

Query Match 1.0%; Score 20; DB 9; Length 642;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 325 AACCAAGCAAGAAATCTTGA 344
DB 74 AACCAAGCAAGAAATCTTGA 93

RESULT 30
LOCUS BG492386 650 bp mRNA linear EST 27-MAR-2001
DEFINITION 602536313F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4655307 5',
ACCESSION BG492386
VERSION BG492386
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 650)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapds@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LCM1444 row: 1 column: 04
High quality sequence stop: 649.
Location/Qualifiers

FEATURES

Source
1. 650
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4655307"
/clone_lib="NIH_MGC_59"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccgcctcgcc); Site_2: SfiI (ggccatcggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATATAGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGAGCGCCGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 258 a 121 c 141 g 129 t 1 others
ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 650;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAAGCAACGAAGCAAGAA 338
DB 274 AAAAGCAACGAAGCAAGAA 293

RESULT 31
LOCUS BB498524 661 bp mRNA linear EST 25-OCT-2001
DEFINITION BB498524 RIKEN full-length enriched, 0 day neonate kidney Mus
musculus cDNA clone D630015B03 3', mRNA sequence.
ACCESSION BB498524
VERSION BB498524
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 661)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,
'M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,
'D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
On Jul 26, 2000 this sequence version replaced gi:9473625.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Saito-cho, Tsukuba-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,
'M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Iwata,M., Ohara,E.,
Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,
'S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,
'Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanae,I., Aizawa,
'K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Location/Qualifiers

FEATURES

Source
1. 661
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="D630015B03"
/clone_lib="RIKEN full-length enriched, 0 day neonate
kidney"
/tissue_type="kidney"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/note="Site_1: Sali; Site_2: BamHI; cDNA library was

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Mammalia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Eukaryota: Eutheria: Rodentia: Sciurognathi; Muridae: Murinae: Mus. 1 (bases 1 to 705)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:357372
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 466.
Location/Qualifiers
1. .705
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:582724"
/clone_1lb="Soares_thymus_2NBM1"
/sex="male"
/tissue_type="thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bernaldo."

BASE COUNT 215 a 163 c 171 g 154 t 2 others
ORIGIN

Query Match 1.0%; Score 20; DB 9; Length 705;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 ACCCAAGCAAGAAATCTTGA 344
|||||
Db 293 ACCCAAGCAAGAAATCTTGA 312

RESULT 35
LOCUS BF658632 710 bp mRNA linear EST 20-DEC-2000
DEFINITION maa98h01.y1 Soares_thymus_2NBM1 Mus musculus cDNA clone
IMAGE:3824881 5' similar to SW:NED1_MOUSE P33215 NEDD1 PROTEIN ;,
RNA sequence.
BF658632
ACCESSION BF658632.1 GI:11923766
VERSION
KEYWORDS house mouse.
SOURCE
ORGANISM Mus musculus
REFERENCE Mammalia: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Eukaryota: Eutheria: Rodentia: Sciurognathi; Muridae: Murinae: Mus. 1 (bases 1 to 710)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@ncl.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1461233
Seq primer: -40RP from Gibco
High quality sequence stop: 462.
Location/Qualifiers
1. .710
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3824881"
/clone_1lb="Soares_thymus_2NBM1"
/sex="male"
/tissue_type="thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bernaldo."

BASE COUNT 221 a 157 c 177 g 155 t
ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 710;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 325 ACCCAAGCAAGAAATCTTGA 344
|||||
Db 647 ACCCAAGCAAGAAATCTTGA 666

RESULT 36
LOCUS AG099717 712 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-102B07.F, genomic survey sequence.
ACCESSION AG099717
VERSION AG099717.1 GI:16720234
KEYWORDS GSS: GSS (genome survey sequence).
SOURCE Pan troglodytes male lymphoblast DNA, clone_1lb:PTB Chimpanzee Male BAC library clone:PTB-102B07.F.
ORGANISM Pan troglodytes
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Pan.
AUTHORS 1 (sites)
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 712)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpes@sc.riken.go.jp, URL:<http://hgp.gsc.riken.go.jp/>, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB this BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS
Sequencing: -21M13
LIBRARY
Vector : PKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .712
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-102807.F"

BASE COUNT 157 a 200 c 158 g 192 t 5 others
ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 712;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 453 CCTTGTGCTGCGAATGCTT 472
|||||
Db 230 CCTTGTGCTGCGAATGCTT 211

RESULT 37
BG506942/c 731 bp mRNA linear EST 27-MAR-2001
LOCUS 601861569F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4071257 5',
DEFINITION mRNA sequence.
ACCESSION BG506942
VERSION BG506942.1 GI:13468459
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 731)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LICM916 row: 1 column: 18
High quality sequence stop: 712.
Location/Qualifiers
1. .731

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4071257"
/clone_1lb="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgccg); Site_2: SfiI (ggcattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCATTATGAGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT 190 a 190 c 133 g 218 t
ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 731;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1439 GACATGATGAAACAGTCTT 1458
|||||
Db 339 GACATGATGAAACAGTCTT 320

RESULT 38
AZ354492 737 bp DNA linear GSS 02-OCT-2000
LOCUS IM0093115R Mouse 10kb plasmid UUGCJM library Mus musculus genomic
DEFINITION clone UUGCJM0093115 R, DNA sequence.
ACCESSION AZ354492
VERSION AZ354492.1 GI:10465987
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 737)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0093 row: 1 column: 15
Seq primer: CACACAGAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 737.
Location/Qualifiers
1. .737

FEATURES
source

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCJM0093115"
/clone_1lb="Mouse 10kb plasmid UUGCJM library"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD22nv, Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114|g14712072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 231 a 178 c 146 g 181 t 1 others

ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 737;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1501 TCAGCGTGTAAACAGCAA 1520
 ||||||||||||||||||
 DB 72 TCAGCGTGTAAACAGCAA 91

RESULT 39

LOCUS BE239667 787 bp mRNA linear EST 14-NOV-2000
 DEFINITION 601906930F1 NIH_MGC_54 Homo sapiens CDNA clone IMAGE:4134678 5',
 mRNA sequence.

ACCESSION BE239667
 VERSION BE239667.1 GI:11153589

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 787)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

FEATURES

Location/Qualifiers

source

1..787

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4134678"

/clone_lib="NIH_MGC_54"

/tissue_type="from chronic myelogenous leukemia"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
 Site.1: SfiI (ggcgctggcc); Site.2: SfiI (ggcgctggcc)
); Double-stranded cDNA was prepared from cell line RNA.
 5' and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGGCCATTATGGC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGCGCGCGGACATG-dt(30)BN-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA)."

BASE COUNT 287 a 146 c 194 g 160 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 787;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAGCAACCAAGCAAGAA 338
 ||||||||||||||||||
 DB 278 AAAGCAACCAAGCAAGAA 297

RESULT 40

BE889059

LOCUS BE889059 788 bp mRNA linear EST 20-OCT-2000
 DEFINITION 60151393F1 NIH_MGC_71 Homo sapiens CDNA clone IMAGE:3915359 5',
 mRNA sequence.

ACCESSION BE889059
 VERSION BE889059.1 GI:10345986
 KEYWORDS EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 788)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

FEATURES

Location/Qualifiers

source

1..788

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3915359"

/clone_lib="NIH_MGC_71"

/tissue_type="telomysarcoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site.1: NotI;
 Site.2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb.
 High quality sequence stop: 716.

BASE COUNT 311 a 138 c 182 g 157 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 788;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAGCAACCAAGCAAGAA 338
 ||||||||||||||||||
 DB 242 AAAGCAACCAAGCAAGAA 261

RESULT 41

LOCUS BE889059 803 bp mRNA linear EST 18-SEP-2001
 DEFINITION 603314371F1 NCI_CGAP_Mam6 Mus musculus CDNA clone IMAGE:534320 5',
 mRNA sequence.

ACCESSION BE889059
 VERSION BE889059.1 GI:15653383
 KEYWORDS EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 803)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM1900 row: n column: 17
High quality sequence stop: 749.

FEATURES

source

1. .803
Location/Qualifiers

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5354320"
/clone_lib="NCL_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"

BASE COUNT

252 a 174 c 182 g 195 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 803;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 325 AACCAACGAGAAATCTGA 344

Db 150 AACCAACGAGAAATCTGA 169

RESULT 42

BF667800 841 bp mRNA linear EST 21-DEC-2000
LOCUS 602122620F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4279709 5',
DEFINITION mRNA sequence.

ACCESSION BF667800
VERSION BF667800.1 GI:11941695
KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 841)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM105 row: 9 column: 06
High quality sequence stop: 534.

FEATURES

source

1. .841
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4279709"
/clone_lib="NIH_MGC_56"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (p1 phage-resistant)"
/note="Organ: Brain; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgcc); Site_2: SfiI (ggcgctcgcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTTGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGCCGACATG-DT(30)BN-3'

(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA).

BASE COUNT 324 a 136 c 220 g 160 t 1 others

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 841;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAGCAACCAAGCAAGAA 338

Db 81 AAAGCAACCAAGCAAGAA 100

RESULT 43

BE383813 861 bp mRNA linear EST 21-JUL-2000
LOCUS 601298315F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3628587 5',
DEFINITION mRNA sequence.

ACCESSION BE383813
VERSION BE383813.1 GI:9329088
KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 861)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM13 row: m column: 04
High quality sequence start: 68
High quality sequence stop: 558.

FEATURES

source

1. .861
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3628587"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 236 a 238 c 214 g 173 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 861;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 GATATCCACAGATTGGT 93

Db 559 GATATCCACAGATTGGT 578

RESULT 44
BH466047
LOCUS BH466047 873 bp DNA Linear GSS 13-DEC-2001
DEFINITION BCGVS80TR BCGV Brassica oleracea genomic clone BCGVS80, DNA sequence.
ACCESSION BH466047 GI:17665680
VERSION BH466047.1
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 873)
AUTHORS Town, C.D., Van Aken, S., Uterback, T. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other_GSSs: BCGVS80TF
Contact: Chris Town
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
Source Location/Qualifiers
1..873
/organism="Brassica oleracea"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BCGVS80"
/clone_lib="BCGV"
/note="Vector: PHOS1, site_1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 linkers"

BASE COUNT 259 a 205 c 207 g 202 t
ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 873;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1643 AACCTGTGACCCACTTGT 1662
|||||
Db 186 AACCGTGTGACCCACTTGT 205

RESULT 45
LOCUS BF670847 888 bp mRNA linear EST 21-DEC-2000
DEFINITION 60214992991 NIH_MGC_81 Homo sapiens CDNA clone IMAGE:4291278 5', mRNA sequence.
ACCESSION BF670847
VERSION BF670847
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 888)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1135 row: 1 column: 07
High quality sequence stop: 622.

FEATURES
Source Location/Qualifiers
1..888
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4291278"
/clone_lib="NIH_MGC_81"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: muscle (skeletal); Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGGCGCATTTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCCGAGGCGCCGACATG-dT(30)-BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 308 a 191 c 197 g 192 t
ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 888;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAACGACCAACCAAGANA 338
|||||
Db 302 AAACGACCAACCAAGANA 321

RESULT 46
LOCUS BF696798 896 bp mRNA linear EST 22-DEC-2000
DEFINITION 602125331P1 NIH_MGC_56 Homo sapiens CDNA clone IMAGE:4282542 5', mRNA sequence.
ACCESSION BF696798
VERSION BF696798
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 896)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1112 row: m column: 07
High quality sequence stop: 563.

FEATURES
Source Location/Qualifiers
1..896
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4282542"
/clone_lib="NIH_MGC_56"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattatggcc); Double-stranded CDNA was prepared from cell line RNA. 5'

and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCGC-3' and 3' adaptor (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 325 a 185 c 207 g 177 t 2 others

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 896;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAGCAACCAAGCAGAGAA 338
|||||

Db 253 AAAGCAACCAAGCAGAGAA 272

RESULT 47
LOCUS BF247101 904 bp mRNA linear EST 14-NOV-2000
DEFINITION 60185454F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4074298 5',
mRNA sequence.
ACCESSION BF247101
VERSION BF247101.1 GI:11162177
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 904)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM924 row: h column: 11
High quality sequence stop: 556.
Location/Qualifiers
1. 904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4074298"
/clone_lib="NIH_MGC_57"
/tissue_type="g10blastoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgctcgcc); Site_2: SfiI (ggcgctatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGCGCGCGCGCATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 320 a 204 c 208 g 172 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 904;
Best Local Similarity 100.0%; Pred. No. 47;

QY 319 AAAGCAACCAAGCAGAGAA 338
|||||

Db 257 AAAGCAACCAAGCAGAGAA 276

RESULT 48
LOCUS BF697265 917 bp mRNA linear EST 22-DEC-2000
DEFINITION 602129793F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:428686 5',
mRNA sequence.
ACCESSION BF697265
VERSION BF697265.1 GI:11982673
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 917)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM123 row: i column: 23
High quality sequence stop: 613.
Location/Qualifiers
1. 917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:428686"
/clone_lib="NIH_MGC_56"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgctcgcc); Site_2: SfiI (ggcgctatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGCGCATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 343 a 187 c 229 g 158 t

ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 917;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAGCAACCAAGCAGAGAA 338
|||||

Db 272 AAAGCAACCAAGCAGAGAA 291

RESULT 49
LOCUS BG282996 938 bp mRNA linear EST 21-FEB-2001
DEFINITION 602406185F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4518260 5',
mRNA sequence.
ACCESSION BG282996
VERSION BG282996.1 GI:13032433

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1AM10412 row: 3 column: 21
 High quality sequence stop: 719.
FEATURES
 Location/Qualifiers
 1..938
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4518260"
 /clone.lib="NIH-MGC_91"
 /tissue_type="adenocarcinoma, cell_line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC library."
BASE COUNT 290 a 152 c 175 g 321 t
ORIGIN

Query Match 1.0%; Score 20; DB 10; Length 938;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1867 TAGTCACTTCACAGCTCA 1886
 ||||||||||||||||||||
 DB 177 TAGTCACTTCACAGCTCA 158

RESULT 50
 AO688268
 LOCUS AO688268 969 bp DNA linear GSS 01-JUL-1999
 DEFINITION nbxb0077M07r CUGI Rice BAC Library Oryza sativa genomic clone
 ACCESSION AO688268
 VERSION AO688268.1 GI:5329352
 KEYWORDS GSS.
 SOURCE Oryza sativa.
 ORGANISM Oryza sativa.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 969)
 Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Rice Genome
 Unpublished (1998)
COMMENT Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: rwing@clemson.edu
 Seq primer: GGAACACCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 338.

FEATURES
 Location/Qualifiers
 1..969
 /organism="Oryza sativa"
 /strain="Japonica"
 /cultivar="Nipponbare"
 /db_xref="taxon:4530"
 /clone="nbxb0077M07r"
 /clone.lib="CUGI Rice BAC Library"
 /tissue_type="leaf"
 /lab_host="E. coli DH10B"
 /note="Vector: pBelBAC11; Site_1: HindIII; Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocytiledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
BASE COUNT 303 a 212 c 173 g 275 t 6 others
ORIGIN

Query Match 1.0%; Score 20; DB 12; Length 969;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 CAAGTAATTACCTGCTTAC 655
 ||||||||||||||||||||
 DB 218 CAAGTAATTACCTGCTTAC 237

Search completed: June 13, 2002, 11:34:17
 Job time: 10484 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 09:37:12 ; Search time 89.1 Seconds
(without alignments)
5510.902 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTACGACAAATATAA.....ATCTTCACACGACACACCA 1999

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size: 15

Total number of hits satisfying chosen parameters: 425

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database:

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1999	100.0	1999	3 US-08-961-083-1	Sequence 1, Appli
2	1438	71.9	2049	3 US-08-481-435-5	Sequence 5, Appli
3	796	39.8	960	2 US-08-245-511-3	Sequence 3, Appli
4	796	39.8	960	2 US-08-600-993A-3	Sequence 3, Appli
5	23	1.2	33	3 US-08-961-083-227	Sequence 227, App
6	19	1.0	2193	1 US-08-731-716-1	Sequence 1, Appli
7	19	1.0	2193	1 US-08-731-716-3	Sequence 3, Appli
8	18	0.9	4136	4 US-09-103-875-2	Sequence 2, Appli
9	17	0.9	27	3 US-08-961-083-228	Sequence 228, App
10	17	0.9	834	4 US-09-020-956-5	Sequence 5, Appli
11	17	0.9	834	4 US-09-030-607-5	Sequence 5, Appli
12	17	0.9	834	4 US-09-439-313-5	Sequence 5, Appli
13	17	0.9	971	4 US-08-630-915A-197	Sequence 197, App
14	17	0.9	1457	4 US-09-444-053-3	Sequence 3, Appli
15	16	0.8	20	4 US-09-444-053-5	Sequence 5, Appli
16	16	0.8	24	4 US-09-457-708-21	Sequence 21, Appli
17	16	0.8	30	1 US-08-232-144-2	Sequence 2, Appli
18	16	0.8	256	2 US-08-230-002-17	Sequence 17, Appli
19	16	0.8	256	2 US-08-678-854-17	Sequence 17, Appli
20	16	0.8	359	2 US-08-687-080-65	Sequence 65, Appli
21	16	0.8	574	1 US-08-592-126-137	Sequence 137, App
22	16	0.8	595	1 US-08-784-289-1	Sequence 1, Appli
23	16	0.8	643	4 US-08-858-207A-116	Sequence 116, App
24	16	0.8	766	4 US-08-858-207A-31	Sequence 31, Appli
25	16	0.8	875	4 US-08-936-165A-125	Sequence 125, App
26	16	0.8	890	1 US-08-592-126-78	Sequence 78, Appli
27	16	0.8	1186	1 US-08-368-236-2	Sequence 2, Appli

C 28	16	0.8	1295	1	US-08-245-295-6	Sequence 6, Appli
C 29	16	0.8	1295	1	US-08-481-130-6	Sequence 6, Appli
C 30	16	0.8	1295	1	US-08-656-984A-6	Sequence 6, Appli
C 31	16	0.8	1295	1	US-08-485-604-6	Sequence 6, Appli
C 32	16	0.8	1295	2	US-08-487-595-6	Sequence 25, Appli
C 33	16	0.8	1295	3	US-08-863-790-25	Sequence 25, Appli
C 34	16	0.8	1295	3	US-08-296-749-25	Sequence 9, Appli
C 35	16	0.8	1314	2	US-08-440-845D-9	Sequence 9, Appli
C 36	16	0.8	1314	3	US-08-868-458-9	Sequence 24, Appli
C 37	16	0.8	1353	2	US-08-611-280-24	Sequence 24, Appli
C 38	16	0.8	1353	4	US-09-195-940-44	Sequence 24, Appli
C 39	16	0.8	1400	2	US-08-001-078A-2	Sequence 2, Appli
C 40	16	0.8	1400	2	US-08-463-218-2	Sequence 2, Appli
C 41	16	0.8	1400	5	PCT-US94-00253-2	Sequence 9, Appli
C 42	16	0.8	1472	1	US-08-245-295-9	Sequence 9, Appli
C 43	16	0.8	1472	1	US-08-481-130-9	Sequence 9, Appli
C 44	16	0.8	1472	1	US-08-656-984A-9	Sequence 9, Appli
C 45	16	0.8	1472	1	US-08-485-604-9	Sequence 9, Appli
C 46	16	0.8	1472	2	US-08-487-595-9	Sequence 9, Appli
C 47	16	0.8	1500	1	US-08-443-568B-15	Sequence 15, Appli
C 48	16	0.8	1500	5	PCT-US94-06997-15	Sequence 15, Appli
C 49	16	0.8	1510	1	US-07-759-568-4	Sequence 4, Appli
C 50	16	0.8	1528	1	US-08-288-839-1	Sequence 4, Appli
C 51	16	0.8	1528	1	US-08-288-839-3	Sequence 3, Appli
C 52	16	0.8	1528	1	US-08-471-206-1	Sequence 1, Appli
C 53	16	0.8	1528	1	US-08-471-206-5	Sequence 5, Appli
C 54	16	0.8	1528	1	US-08-281-025D-1	Sequence 1, Appli
C 55	16	0.8	1547	1	US-08-288-839-4	Sequence 4, Appli
C 56	16	0.8	1547	1	US-08-471-206-8	Sequence 8, Appli
C 57	16	0.8	1568	4	US-08-884-077-3	Sequence 3, Appli
C 58	16	0.8	1605	6	5244676-1	Sequence 77, Appli
C 59	16	0.8	1639	2	US-08-468-819-77	Sequence 8, Appli
C 60	16	0.8	1727	6	US-08-202-056-8	Sequence 29, Appli
C 61	16	0.8	1748	1	US-07-912-900-29	Sequence 29, Appli
C 62	16	0.8	1755	1	US-08-285-309-29	Sequence 29, Appli
C 63	16	0.8	1755	2	US-08-502-046-29	Sequence 29, Appli
C 64	16	0.8	1755	1	US-08-502-046-29	Sequence 29, Appli
C 65	16	0.8	1801	1	US-08-557-917A-1	Sequence 1, Appli
C 66	16	0.8	1801	3	US-09-084-153-1	Sequence 1, Appli
C 67	16	0.8	1801	3	US-09-084-079-1	Sequence 1, Appli
C 68	16	0.8	1812	1	US-07-912-900-28	Sequence 28, Appli
C 69	16	0.8	1812	1	US-08-285-309-28	Sequence 28, Appli
C 70	16	0.8	1812	1	US-08-313-075A-37	Sequence 37, Appli
C 71	16	0.8	1812	2	US-08-502-046-28	Sequence 28, Appli
C 72	16	0.8	1824	3	US-08-606-505B-1	Sequence 1, Appli
C 73	16	0.8	1824	4	US-09-616-990-1	Sequence 1, Appli
C 74	16	0.8	1931	4	US-08-868-435-28	Sequence 28, Appli
C 75	16	0.8	1931	4	US-08-744-231-28	Sequence 28, Appli
C 76	16	0.8	2018	3	US-08-714-918-16	Sequence 16, Appli
C 77	16	0.8	2018	4	US-09-265-315-16	Sequence 16, Appli
C 78	16	0.8	2018	4	US-09-265-315-16	Sequence 16, Appli
C 79	16	0.8	2018	4	US-09-266-417-16	Sequence 16, Appli
C 80	16	0.8	2088	1	US-08-332-838-1	Sequence 1, Appli
C 81	16	0.8	2088	4	US-09-458-481B-3	Sequence 3, Appli
C 82	16	0.8	2214	1	US-08-245-295-7	Sequence 7, Appli
C 83	16	0.8	2214	1	US-08-481-130-7	Sequence 7, Appli
C 84	16	0.8	2214	1	US-08-656-984A-7	Sequence 7, Appli
C 85	16	0.8	2214	1	US-08-485-604-7	Sequence 7, Appli
C 86	16	0.8	2214	2	US-08-487-595-7	Sequence 7, Appli
C 87	16	0.8	2268	3	US-09-344-579-1	Sequence 1, Appli
C 88	16	0.8	2268	3	US-08-714-918-2	Sequence 2, Appli
C 89	16	0.8	2368	4	US-09-265-315-2	Sequence 2, Appli
C 90	16	0.8	2368	4	US-09-265-315-2	Sequence 2, Appli
C 91	16	0.8	2368	4	US-09-266-417-2	Sequence 2, Appli
C 92	16	0.8	2550	1	US-08-245-295-10	Sequence 10, Appli
C 93	16	0.8	2550	1	US-08-481-130-10	Sequence 10, Appli
C 94	16	0.8	2550	1	US-08-656-984A-10	Sequence 10, Appli
C 95	16	0.8	2550	1	US-08-485-604-10	Sequence 10, Appli
C 96	16	0.8	2550	2	US-08-487-595-10	Sequence 10, Appli
C 97	16	0.8	2936	4	US-08-943-731-64	Sequence 64, Appli
C 98	16	0.8	2988	1	US-08-245-295-1	Sequence 1, Appli
C 99	16	0.8	2988	1	US-08-481-130-1	Sequence 1, Appli
C 100	16	0.8	2988	1	US-08-656-984A-1	Sequence 1, Appli


```

c 393 15 0.8 9606 2 US-08-370-235A-1 Sequence 1, Appl1
c 394 15 0.8 9687 4 US-09-133-944-2 Sequence 2, Appl1
c 395 15 0.8 9919 3 US-08-880-179-1 Sequence 1, Appl1
c 396 15 0.8 10322 4 US-09-330-330-3 Sequence 3, Appl1
c 397 15 0.8 10348 2 US-08-457-273B-41 Sequence 41, Appl1
c 398 15 0.8 10348 3 US-08-556-419-13 Sequence 13, Appl1
c 399 15 0.8 10348 4 US-09-041-886-14 Sequence 14, Appl1
c 400 15 0.8 10366 1 US-08-246-982A-5 Sequence 5, Appl1
c 401 15 0.8 10366 1 US-08-453-265-5 Sequence 5, Appl1
c 402 15 0.8 10803 3 US-09-080-044-1 Sequence 1, Appl1
c 403 15 0.8 11907 4 US-08-061-376-4 Sequence 4, Appl1
c 404 15 0.8 12847 1 US-08-550-715-1 Sequence 1, Appl1
c 405 15 0.8 14255 1 US-08-320-559-1 Sequence 1, Appl1
c 406 15 0.8 14255 1 US-08-327-392-1 Sequence 1, Appl1
c 407 15 0.8 14255 1 US-08-306-691B-55 Sequence 55, Appl1
c 408 15 0.8 14255 5 PCT-US94-04496-1 Sequence 1, Appl1
c 409 15 0.8 14255 5 PCT-US94-04496-1 Sequence 1, Appl1
c 410 15 0.8 14507 3 US-08-785-150-1 Sequence 1, Appl1
c 411 15 0.8 14507 4 US-09-660-299-1 Sequence 1, Appl1
c 412 15 0.8 14507 4 US-09-435-377-1 Sequence 1, Appl1
c 413 15 0.8 19227 3 US-09-090-793-13 Sequence 13, Appl1
c 414 15 0.8 20084 4 US-08-943-731-5 Sequence 5, Appl1
c 415 15 0.8 22846 2 US-08-469-461-3 Sequence 3, Appl1
c 416 15 0.8 22846 3 US-07-890-609-3 Sequence 3, Appl1
c 417 15 0.8 28994 3 US-08-884-324-14 Sequence 14, Appl1
c 418 15 0.8 35100 1 US-08-306-691B-19 Sequence 19, Appl1
c 419 15 0.8 35100 5 PCT-US93-06251-19 Sequence 19, Appl1
c 420 15 0.8 40138 3 US-09-090-793-12 Sequence 12, Appl1
c 421 15 0.8 48974 4 US-08-920-422-17 Sequence 17, Appl1
c 422 15 0.8 51259 3 US-08-781-891-209 Sequence 209, App
c 423 15 0.8 51952 3 US-08-947-823-1 Sequence 1, Appl1
c 424 15 0.8 90050 4 US-09-245-041-5 Sequence 5, Appl1
c 425 15 0.8 4411529 4 US-09-103-840A-1 Sequence 1, Appl1

```

ALIGNMENTS

```

RESULT 1
US-08-961-083-1
: Sequence 1, Application US/08961083
: Patent No. 6159469
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 1:

```

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1999 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
US-08-961-083-1

Query Match      100.0%; Score 1999; DB 3; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTACGACAAATAAATCACTATGCTGACTGGTTCGTGACGCCGCTCAA 60
DB 1 TAAATCTACGACAAATAAATCACTATGCTGACTGGTTCGTGACGCCGCTCAA 60

QY 61 TGCCCAAGCTAATGATATTCGCCACAGTTTGGTTAAGGCATGTTCTATCCAGACCA 120
DB 61 TGCCCAAGCTAATGATATTCGCCACAGTTTGGTTAAGGCATGTTCTATCCAGACCA 120

QY 121 TCGCTTCTTGACCCACAGGGGGATTGATACATCCGATCCTGGAGCTTTCTTGCGCAA 180
DB 121 TCGCTTCTTGACCCACAGGGGGATTGATACATCCGATCCTGGAGCTTTCTTGCGCAA 180

QY 181 TCTGCAAAAGCAATTCCTCCAGAGTGATCACTCTCCCAACAGTTGATTAAGTTGAC 240
DB 181 TCTGCAAAAGCAATTCCTCCAGAGTGATCACTCTCCCAACAGTTGATTAAGTTGAC 240

QY 241 TTACTTTTCACTTCGACCTTCGACCACTATTTTCGTAAGGCTCAGGAACCTGGTT 300
DB 241 TTACTTTTCACTTCGACCTTCGACCACTATTTTCGTAAGGCTCAGGAACCTGGTT 300

QY 301 AGCGATTCAGTGTGAACAAACCAACCAAGCAAGAAATCTGACCTACTATTAATAA 360
DB 301 AGCGATTCAGTGTGAACAAACCAACCAAGCAAGAAATCTGACCTACTATTAATAA 360

QY 361 GGTCTACATCTCTAATGGAGCACTATGATGACAGACGCTCAAAACTATATGGTAA 420
DB 361 GGTCTACATCTCTAATGGAGCACTATGATGACAGACGCTCAAAACTATATGGTAA 420

QY 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGGCTCAGGCACC 480
DB 421 AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGGCTGGGAATGGCTCAGGCACC 480

QY 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCAGCCGGAATCTGGTCTT 540
DB 481 AAACCAATATGACCCCTATTCACATCCAGAGCAGCCAGCCGGAATCTGGTCTT 540

QY 541 ATCTGAATGAATAAATCAAGGCTACATCTCTGCTGAACAGTATGANAAGCACTAATAC 600
DB 541 ATCTGAATGAATAAATCAAGGCTACATCTCTGCTGAACAGTATGANAAGCACTAATAC 600

QY 601 ACCAATTAAGTATGATCAAGAGTCAATAGCAAGTATTAATCCCTGCTACATGGA 660
DB 601 ACCAATTAAGTATGATCAAGAGTCAATAGCAAGTATTAATCCCTGCTACATGGA 660

QY 661 TAAATTAACCTCAAGAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTAATACCTACTCAC 720
DB 661 TAAATTAACCTCAAGAGTCAATCAATCAAGTTGAAGAAGAAACAGGCTAATACCTACTCAC 720

QY 721 AACTGGAGTATGATCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGGGGATAT 780
DB 721 AACTGGAGTATGATCTACCAAAATGTAGACCAAGAGCTCAAAAACATCTGGGGATAT 780

QY 781 TTACAAATACGAGCAATACCTGCTATCCAGACGATGAATTCGAAGTCCCTCTACCAT 840
DB 781 TTACAAATACGAGCAATACCTGCTATCCAGACGATGAATTCGAAGTCCCTCTACCAT 840

QY 841 TGTGATGTTTCTAACGGTAAATGATTCGCGAGCTAGAGCAGCGCATGACGTAAGTAA 900
DB 841 TGTGATGTTTCTAACGGTAAATGATTCGCGAGCTAGAGCAGCGCATGACGTAAGTAA 900

QY 901 TGTTCTCTCGAATTAACCAAGCAGTAGAAGCAACCGGAGCTGGGATCAACTATGAA 960

```


D	b	108	TGCCAACGCTAATATATATTTCCACAGATTTTGGTTAAGCAATCGTTTCTATCGAAGACCA	167
O	y	121	TGCGTCTTGACACACAGGGGGATGTGATACATCGGTATCTGGAGCTTTCTTGCGCAA	180
D	b	168	TGCGTCTTTCACACACAGGGGGATGTGATACATCGGTATCTGGAGCTTTCTTGCGCAA	227
O	y	181	TCTGCAAGCAATTCCTCCACAGGTGGATCACTCCACCAAGTGTGATTAAGTTAC	240
D	b	228	TCTGCAAGCAATTCCTCCACAGGTGGATCACTCCACCAAGTGTGATTAAGTTAC	287
O	y	241	TTACATTTCAACTTCGACTTCGCGACGACATATTTCTGTAAGGCTCGAGAGCTTGCTT	300
D	b	288	TTACTTTTCAACTTCGACTTCGCGACGACATATTTCTGTAAGGCTCGAGAGCTTGCTT	347
O	y	301	AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	360
D	b	348	AGCGATTCAGTTAGAACAAAAAGCAACCAAGCAAGAAATCTTGACCTACTATATAATAA	407
O	y	361	GGTCTACATGCTATATGGAACTATGGAAATGACAGACAGCTCCAAACTCTATATGGTAA	420
D	b	408	GGTCTACATGCTATATGGAACTATGGAAATGACAGACAGCTCCAAACTACTATATGGTAA	467
O	y	421	AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGCTGGTGGAAATGCTCAGACGACC	480
D	b	468	AGACCTCAATTAATTAAGTTTACCTCAGTTAGCTTGCTGGTGGAAATGCTCAGACGACC	527
O	y	481	AAACCAATATGACCCCTATTCACATCCAGACGACCCCAAGACCGCGAAACTTGCTT	540
D	b	528	AAACCAATATGACCCCTATTCACATCCAGACGACCCCAAGACCGCGAAACTTGCTT	587
O	y	541	ATTCGAATGAAAAATCAAGGTCATCTCTGCTGAACGATGTGAGAAAGCAAGTCATAAC	600
D	b	588	ATTCGAATGAAAAATCAAGGTCATCTCTGCTGAACGATGTGAGAAAGCAAGTCATAAC	647
O	y	601	ACCAATTTACTATGACTACAGAAAGTCTCAATGACAAATTAATTAATCCGTCTACATGGA	660
D	b	648	ACCAATTTACTATGACTACAGAAAGTCTCAATGACAAATTAATTAATCCGTCTACATGGA	707
O	y	661	TAAATTAACCTCAGGAAGTCAATCAATCAAGTTGAGAGAAACAGGCTATTAACCTACTAC	720
D	b	708	TAAATTAACCTCAGGAAGTCAATCAATCAAGTTGAGAGAAACAGGCTATTAACCTACTAC	767
O	y	721	AACGGGATGGAATGCTACACAAATGTAAGCAAGAGCTCAAAAACATCTGTGGGATAT	780
D	b	768	AACGGGATGGAATGCTACACAAATGTAAGCAAGAGCTCAAAAACATCTGTGGGATAT	827
O	y	781	TTACATTTACAGACAAATACGTTGCTCCTATCCAGACGATGGAATGCAAGTGCCTTACCAT	840
D	b	828	TTACATTTACAGACAAATACGTTGCTCCTATCCAGACGATGGAATGCAAGTGCCTTACCAT	887
O	y	841	TGTTGATGTTTCTAAGGTTAAAGTCATTTGCCACGCTAGGAGACGCCATCACTCAAGTAA	900
D	b	888	TGTTGATGTTTCTAAGGTTAAAGTCATTTGCCACGCTAGGAGACGCCATCACTCAAGTAA	947
O	y	901	TGTTTCCCTTGCGAATTTAACCAAGCAGTGAAGAAACAAACCCGCACTGGGGATCAACTATGAA	960
D	b	948	TGTTTCCCTTGCGAATTTAACCAAGCAGTGAAGAAACAAACCCGCACTGGGGATCAACTATGAA	1007
O	y	961	ACCGATCAGACATATCTCTGCTGCTGGAGTACGGGTGCTACGATTCATACTGCTACAT	1020
D	b	1008	ACCGATCAGACATATCTCTGCTGCTGGAGTACGGGTGCTACGATTCATACTGCTACAT	1067
O	y	1021	CGTTACAGATGAGCCCTTAATTAACCTCGGGACAAATTAATCTCTGTTTAACTGGGATAG	1080
D	b	1068	CGTTACAGATGAGCCCTTAATTAACCTCGGGACAAATTAATCTCTGTTTAACTGGGATAG	1127
O	y	1081	GGGCTACTTTGGCAACATCACCTTGCAATACGCTTCGCAACAATGCGGAAGCTCCAGC	1140
D	b	1128	GGGCTACTTTGGCAACATCACCTTGCAATACGCTTCGCAACAATGCGGAAGCTCCAGC	1187
O	y	1141	CGTGGAAACCTTAACAAAGGTGGGACGTCAACCGCGCAAGACTTTTCTAAAGGTCTAGG	1200
D	b	1188	CGTGGAAACCTTAACAAAGGTGGGACGTCAACCGCGCAAGACTTTTCTAAAGGTCTAGG	1247

QY	1201	AATGACCTACCCAGTATTTTCATCACTCAATTCGCAATTTTCAAGTAACACACACGCAATTCGA	1266
Db	1248	AATGACCTACCCAGTATTTTCATCACTCAATTCGCAATTTTCAAGTAACACACACGCAATTCGA	1307
QY	1261	CAAAAAAATATGGACCAAGTAGTGAAGAAAGATGGCTGGCTGTTAAGCGCTGCTTTGGCAAAATGG	1320
Db	1308	CAAAAAAATATGGACCAAGTAGTGAAGAAAGATGGCTGGCTGTTAAGCGCTGCTTTGGCAAAATGG	1367
QY	1321	TGGAACTTACTATTAACCAATGTATATCCATAAAGTCGCTTTAGTAGTAGGAGTGA AAA	1380
Db	1368	TGGAACTTACTATTAACCAATGTATATCCATAAAGTCGCTTTAGTAGTAGGAGTGA AAA	1427
QY	1381	AGAGTCTCTTAATGCGAACTGCTGGCAATGAAGAAAGACAGCCTATATATGATACCGA	1440
Db	1428	AGAGTCTCTTAATGCGAACTGCTGGCAATGAAGAAAGACAGCCTATATATGATACCGA	1487
QY	1441	CATGATGAAAAACAGTCTTGACTTATGGAACGTGAGCAAAAAATGCCATCTTCTTGGCTGCC	1500
Db	1488	CATGATGAAAAACAGTCTTGAAGTTATGGAACGTGAGCAAAAAATGCCATCTTCTTGGCTGCC	1547
QY	1501	TCAGGCTGTGTA AAAACAGGAACCTCTAATCTTAACAAGACGAGAAATTTGAAAACCAATCA	1566
Db	1548	TCAGGCTGTGTA AAAACAGGAACCTCTAATCTTAACAAGACGAGAAATTTGAAAACCAATCA	1607
QY	1561	GACCTCTCAATTTTGTAGCACCTGATGAACATATTTGCTGGCTATACCGCTAATATTCAAT	1620
Db	1608	GACCTCTCAATTTTGTAGCACCTGATGAACATATTTGCTGGCTATACCGCTAATATTCAAT	1667
QY	1621	GCGTGTATGACAGGCTATTTCTTAACCGCTGACACACATTTGTAAGGCATGGCCTTAGCGT	1680
Db	1668	GCGTGTATGACAGGCTATTTCTTAACCGCTGACACACATTTGTAAGGCATGGCCTTAGCGT	1727
QY	1681	CGCTGCCAAATTTTACCGGCTCTAATGATGACCTACCTGCTGGAAGGAAGCAATCCGAAAGA	1740
Db	1728	CGCTGCCAAATTTTACCGGCTCTAATGATGACCTACCTGCTGGAAGGAAGCAATCCGAAAGA	1787
QY	1741	TTGGAATTTACCAAGGAGGCTCTACAGAAATGGAGAATTCGATTTTAAAAATGGTGTCTCG	1800
Db	1788	TTGGAATTTACCAAGGAGGCTCTACAGAAATGGAGAATTCGATTTTAAAAATGGTGTCTCG	1847
QY	1801	TTTCTACGTGGAACCTACCTGCTGCACACAACCCCATCATACTGAAGATTTCAAGCTCATC	1866
Db	1848	TTTCTACGTGGAACCTACCTGCTGCACACAACCCCATCATACTGAAGATTTCAAGCTCATC	1907
QY	1861	ATCGAATGTCAACTTCAACAGTCTAGCTCAACCATCCATCAAGCAACAATAATATAGTACAC	1920
Db	1908	ATCGAATGTCAACTTCAACAGTCTAGCTCAACCATCCATCAAGCAACAATAATATAGTACAC	1967
QY	1921	TACCAATCTTAACAAATATATAGCAACAAATCAATTAACMACCCCTGATCAACAAATCAGAA	1980
Db	1968	TACCAATCTTAACAAATATATAGCAACAAATCAATTAACMACCCCTGATCAACAAATCAGAA	2027
QY	1981	TCCTTCACACGACACACCA 1999	
Db	2028	TCCTTCACACGACACACCA 2046	
RESULT 3			
US-08-245-511-3			
; Sequence 3, Application US/08245511			
; Patent No. 5928900			
; GENERAL INFORMATION:			
; APPLICANT: Measure, H Robert			
; APPLICANT: Pearce, Barbara J			
; APPLICANT: Tuomanen, Elaine			
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND			
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON			
; NUMBER OF SEQUENCES: 58			
; CORRESPONDENCE ADDRES:			
; ADDRESSEE: Klauber & Jackson			
; STREET: 411 Hackensack Avenue			
; City: Hackensack			

RESULT 3
US-08-245-511-3
; Sequence 3, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomaten, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klaber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack


```
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 960 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: R6
IMMEDIATE SOURCE:
CLONE: SPR042
FEATURE:
NAME/KEY: CDS
LOCATION: 1..960
US-08-245-511-3

Query Match          39.8%; Score 796; DB 2; Length 960;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TAAATCTACGACAAATCAATCAATCTGCTGCTGGTCTGACCGCGCTCAA 60
DB 12 TAAATCTACGACAAATCAATCAATCTGCTGCTGGTCTGACCGCGCTCAA 71
OY 61 TGGCCAACTATGATATTCACAGATTTGGTTAAAGCAATCTTCTATCGAAGACCA 120
DB 72 TGGCCAACTATGATATTCACAGATTTGGTTAAAGCAATCTTCTATCGAAGACCA 131
OY 121 TCGCTTCTGACACAGAGGGGATGATGATCCGATCCGTGAGAGCTTCTTCCGCAA 180
DB 132 TCGCTTCTGACACAGAGGGGATGATGATCCGATCCGTGAGAGCTTCTTCCGCAA 191
OY 181 TGTGAAAGCAATTCCTCCAGAGTGATCAACTCCACCAAGTTGATTAGTTGAC 240
DB 192 TGTGAAAGCAATTCCTCCAGAGTGATCAACTCCACCAAGTTGATTAGTTGAC 251
OY 241 TTACTTTCACTGCTCGACACAGATTTCTGTAAGGCTCAGGAAGCTTGCTT 300
DB 252 TTACTTTCACTGCTCGACACAGATTTCTGTAAGGCTCAGGAAGCTTGCTT 311
OY 301 AGCGATTGATTAGAAACAAAGCAACAGCAAGAAATCTTGACCTATATATAATA 360
DB 312 AGCGATTGATTAGAAACAAAGCAACAGCAAGAAATCTTGACCTATATATAATA 371
OY 361 GGTCTACATGCTATATGGAACTATGGAATGACAGACAGCTCAAACTACTATGTTAA 420
DB 420 GGTCTACATGCTATATGGAACTATGGAATGACAGACAGCTCAAACTACTATGTTAA 420
```

```
DB 372 GGTCTACATGCTATATGGAACTATGGAATGACAGACAGCTCAAACTACTATGTTAA 431
OY 421 AGACCTCAATATATTAAGTTTACCTGCTAGTTAGCTTGGCTGGAAATGCTCAGGACCC 480
DB 432 AGACCTCAATATATTAAGTTTACCTGCTAGTTAGCTTGGCTGGAAATGCTCAGGACCC 491
OY 481 AAACCAATATGACCCCTATTTACATCCAGACAGCCCAAGACCCGCAAACTTGCTTT 540
DB 492 AAACCAATATGACCCCTATTTACATCCAGACAGCCCAAGACCCGCAAACTTGCTTT 551
OY 541 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCATATAC 600
DB 552 ATCTGAAATGAAAAATCAAGGCTACATCTCTGCTGACAGTATGAGAAAGCATATAC 611
OY 601 ACCAATTCGTGATGACATCAAGTCTCAATACAGCAATATTTCCCTGCTTACATGGA 660
DB 612 ACCAATTCGTGATGACATCAAGTCTCAATACAGCAATATTTCCCTGCTTACATGGA 671
OY 661 TAATTACCTCAAGAAAGTCAATCAAGTGAAGAAAGCAAGCTATTAACCTACTCAC 720
DB 672 TAATTACCTCAAGAAAGTCAATCAAGTGAAGAAAGCAAGCTATTAACCTACTCAC 731
OY 721 AACTGGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTTGGAAT 780
DB 732 AACTGGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTTGGAAT 791
OY 781 TTACAATACAGCAAGTATGCTGCTATCCAGACATGAAATGCAAGTCTTACCAT 840
DB 792 TTACAATACAGCAAGTATGCTGCTATCCAGACATGAAATGCAAGTCTTACCAT 851
OY 841 TGTGATGCTTTTAAACGGTAAGTATGCTGCTATCCAGTACAGCAGCCATCACTCAAGTAA 900
DB 852 TGTGATGCTTTTAAACGGTAAGTATGCTGCTATCCAGTACAGCAGCCATCACTCAAGTAA 911
OY 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAAACAAACCGCAGCTGGGGA 949
DB 912 TGTTCCTTGGGAATTAACCAAGCAGTAGAAACAAACCGCAGCTGGGGA 960

RESULT 4
US-08-600-993A-3
Sequence 3, Application US/08600993A
Patent No. 5981229
GENERAL INFORMATION:
APPLICANT: Measure, H Robert
APPLICANT: Pearce, Barbara J
APPLICANT: Tuomenen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,993A
FILING DATE: 1-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 435
APPLICATION DATA:
APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994
```

ATTORNEY/AGENT INFORMATION:
 NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 600-1-069 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201 487-5800
 TELEFAX: 201 343-1684
 TELEX: 133521
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 960 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Streptococcus pneumoniae
 STRAIN: R6
 IMMEDIATE SOURCE:
 CLONE: SPRU42
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..960
 US-08-961-083-1

Query Match 39.8%; Score 796; DB 2; Length 960;

Best Local Similarity 99.7%; Pred. No. 0;
 Matches 946; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 TAAATCTACGACAAATTAATCACTATTGCTGACCTGGTTGCAAGCCGCGCAA 60
 12 TAAATCTACGACAAATTAATCACTATTGCTGACCTGGTTGCAAGCCGCGCAA 71
 61 TGGCCAGCTAATGATATTCCTCCACAGATTTGTTAAGGCATGTTCTATGCAAGCA 120
 72 TGGCCAGCTAATGATATTCCTCCACAGATTTGTTAAGGCATGTTCTATGCAAGCA 131
 121 TGGCTCTCTGACACAGGGGGATGATNCCATCCGATCCCTGGAGCTTCTTGCGCAA 180
 132 TGGCTCTCTGACACAGGGGGATGATNCCATCCGATCCCTGGAGCTTCTTGCGCAA 191
 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCCACCAAGTGTGATTAAGTTGAC 240
 192 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCCACCAAGTGTGATTAAGTTGAC 251
 241 TTTACTTTTCACTTGCACCTTCCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGCTT 300
 252 TTTACTTTTCACTTGCACCTTCCGACAGACTATTTCTGTAAGGCTCAGGAAGCTTGCTT 311
 301 AGGATTCAGTTGAACAAAGCAACCAAGCAAGAAATCTGACCTACTATTAATAA 360
 312 AGGATTCAGTTGAACAAAGCAACCAAGCAAGAAATCTGACCTACTATTAATAA 371
 361 GGTCTACATGCTAATGGAAGCATGGAATGACAGACAGAGCTCAAAAACACTATGTTAA 420
 372 GGTCTACATGCTAATGGAAGCATGGAATGACAGACAGAGCTCAAAAACACTATGTTAA 431
 421 AGACCTCAATTAATTTAAGTTTACCTAGTTAGCTTGGTGGTGAATGGCTCAGGAC 480
 432 AGACCTCAATTAATTTAAGTTTACCTAGTTAGCTTGGTGGTGAATGGCTCAGGAC 491
 481 AAACCAATATGACCCCTATTACATCCAGAGAGAGCCGCGGAACTTGGTCTT 540
 492 AAACCAATATGACCCCTATTACATCCAGAGAGAGCCGCGGAACTTGGTCTT 551
 541 ATCTGAAATGAAGAAATCAAGGCTACATCTGCTGTAAGCAATGAAGAAACACTCAATAC 600
 552 ATCTGAAATGAAGAAATCAAGGCTACATCTGCTGTAAGCAATGAAGAAACACTCAATAC 611
 601 ACCAATTACTGATGACTACAAAGTCTCAATCAGCAAGTAATACCTTACTACATGA 660

|||||
 Db 612 ACCAATTACTGATGGGCTACAAAGTCTCAATCAGCAAGTAATACCTTACTACATGA 671
 661 TAATTACTTAAGGAAGTATCATCAATGATTAAGAAAGAAACAGGCTATTAACCTACTAC 720
 672 TAATTACTTAAGGAAGTATCATCAATGATTAAGAAAGAAACAGGCTATTAACCTACTAC 731
 721 AACTGGATGATGCTGTACACAAATGATAGACAAAGAAAGTCAAAAACATCTGTGGATAT 780
 732 AACTGGATGATGCTGTACACAAATGATAGACAAAGAAAGTCAAAAACATCTGTGGATAT 791
 781 TTACAATACAGACGAATAGCTGCTATCCAGACGATGAATTCGAAGTCTTACCAT 840
 792 TTACAATACAGACGAATAGCTGCTATCCAGACGATGAATTCGAAGTCTTACCAT 851
 841 TGTGATGTTTCTACGCTTAAGTCAATGCTGCGCAGCTAGAGACGCCATGCTCAAGTAA 900
 852 TGTGATGTTTCTACGCTTAAGTCAATGCTGCGCAGCTAGAGACGCCATGCTCAAGTAA 911
 901 TGTTCCTTGGGAATTAACCAAGCAGTAGAAGAAACCGGAGCTGGGGA 949
 912 TGTTCCTTGGGAATTAACCAAGCAGTAGAAGAAACCGGAGCTGGGGA 960

RESULT 5
 US-08-961-083-227
 ; Sequence 227, Application US/08961083
 ; Patent No. 6159469

GENERAL INFORMATION:
 APPLICANT: Choi et. al.
 TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
 NUMBER OF SEQUENCES: 452
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/961,083
 FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 227:

SEQUENCE CHARACTERISTICS:

LENGTH: 33 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-083-227

Query Match 1.2%; Score 23; DB 3; Length 33;
 Best Local Similarity 100.0%; Pred. No. 0.033;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TAAATCTACGACAAATTAATCACTATTGCTGACCTGGTTGCAAGCCGCGCAA 23
 |||||||

Db 11 TAAATCTACGACATATAAATC 33

RESULT 6

US-08-731-716-1

Sequence 1, Application US/08731716

Patent No. 5789202

GENERAL INFORMATION:

APPLICANT: Hoskins, Joann

APPLICANT: Jaskunas, S. Richard

APPLICANT: Rostock, Pamela K.

APPLICANT: Zhao, Genshi

APPLICANT: Rostock, Paul R. Jr.

APPLICANT: No. 5789202r/s, Franklin H.

TITLE OF INVENTION: Penicillin Binding Protein From

TITLE OF INVENTION: Streptococcus Pneumoniae

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/731,716

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: X-10,887

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2193 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHEICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2193

US-08-731-716-1

Query Match 1.0%; Score 19; DB 1; Length 2193;

Best Local Similarity 100.0%; Pred. No. 4.4;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 TACGCTGCCTTGCAATG 1319

|||||

Db 1555 TACGCTGCCTTGCAATG 1573

RESULT 7

US-08-731-716-3

Sequence 3, Application US/08731716

Patent No. 5789202

GENERAL INFORMATION:

APPLICANT: Hoskins, Joann

APPLICANT: Jaskunas, S. Richard

APPLICANT: Rostock, Pamela K.

APPLICANT: Zhao, Genshi

APPLICANT: Rostock, Paul R. Jr.

APPLICANT: No. 5789202r/s, Franklin H.

TITLE OF INVENTION: Penicillin Binding Protein From

TITLE OF INVENTION: Streptococcus Pneumoniae

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Eli Lilly and Company

STREET: Lilly Corporate Center

CITY: Indianapolis

STATE: Indiana

COUNTRY: U.S.

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/731,716

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Webster, Thomas D.

REGISTRATION NUMBER: 39,872

REFERENCE/DOCKET NUMBER: X-10,887

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 2193 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: mRNA

HYPOTHEICAL: NO

ANTI-SENSE: NO

US-08-731-716-3

Query Match 1.0%; Score 19; DB 1; Length 2193;

Best Local Similarity 68.4%; Pred. No. 4.4;

Matches 13; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1301 TACGCTGCCTTGCAATG 1319

|||||

Db 1555 UACGCGCCUUGCAAAUG 1573

RESULT 8

US-09-103-875-2/c

Sequence 2, Application US/09103875A

Patent No. 6221849

GENERAL INFORMATION:

APPLICANT: Szyl, Moshe

APPLICANT: Bigey, Pascal

APPLICANT: Ramchandani, Shyam

TITLE OF INVENTION: DNA METHYLTRANSFERASE GENOMIC SEQUENCES AND ANTISENSE

TITLE OF INVENTION: OLIGONUCLEOTIDES

FILE REFERENCE: 106101.194

CURRENT APPLICATION NUMBER: US/09/103,875A

CURRENT FILING DATE: 1998-06-24

EARLIER APPLICATION NUMBER: 60/069,865

EARLIER FILING DATE: 1997-12-17

EARLIER APPLICATION NUMBER: 08/866,340

EARLIER FILING DATE: 1997-05-30

NUMBER OF SEQ ID NOS: 138

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2

LENGTH: 4136

TYPE: DNA

ORGANISM: Homo sapiens

US-09-103-875-2

Query Match 0.9%; Score 18; DB 4; Length 4136;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1589 CTATTGCTGGCTATACG 1606
DB 3382 CTATTGCTGGCTATACG 3365

RESULT 9

US-08-961-083-228/C
; Sequence 228, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MD storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 228:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-228

Query Match 0.9%; Score 17; DB 3; Length 27;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1983 CTCACACGACACACCA 1999
DB 27 CTCACACGACACACCA 11

RESULT 10

US-09-020-956-5/C
; Sequence 5, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillin, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-020-956-5

Query Match 0.9%; Score 17; DB 4; Length 834;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AATTAATCACTCAT 30
DB 516 AATTAATCACTCAT 500

RESULT 11

US-09-030-607-5/C
; Sequence 5, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillin, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

LENGTH: 834 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-030-607-5

Query Match 0.9%; Score 17; DB 4; Length 834;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AATAAAATCAACTCAT 30
|||||
DB 516 AATAAAATCAACTCAT 500

RESULT 12
US-09-439-313-5/C
Sequence 5, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiaqun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqi
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retler, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 834
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(834)
OTHER INFORMATION: n = A,T,C or G
US-09-439-313-5

Query Match 0.9%; Score 17; DB 4; Length 834;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AATAAAATCAACTCAT 30
|||||
DB 516 AATAAAATCAACTCAT 500

RESULT 13
US-08-630-915A-197/C
Sequence 197, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
TITLE OF INVENTION: USING SAME

NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 197:
SEQUENCE CHARACTERISTICS:
LENGTH: 971 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-630-915A-197

Query Match 0.9%; Score 17; DB 4; Length 971;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 AGGCTCAGGAGCTTGG 298
|||||
DB 373 AGGCTCAGGAGCTTGG 357

RESULT 14
US-09-444-053-3/C
Sequence 3, Application US/09444053A
Patent No. 6165728
GENERAL INFORMATION:
APPLICANT: Donna T. Ward
APPLICANT: Lex M. Cowser
TITLE OF INVENTION: ANTISENSE MODULATION OF NCK-2 EXPRESSION
FILE REFERENCE: RTS-0122
CURRENT APPLICATION NUMBER: US/09/444,053A
CURRENT FILING DATE: 1999-11-19
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 3
LENGTH: 1457
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (133)..(1275)
US-09-444-053-3

Query Match 0.9%; Score 17; DB 4; Length 1457;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 AGGCTCAGGAGCTTGG 298
|||||
DB 677 AGGCTCAGGAGCTTGG 661

RESULT 15
US-09-444-053-55
; Sequence 53, Application US/09444053A
; Patent No. 6165728
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF NCK-2 EXPRESSION
; FILE REFERENCE: RTS-0122
; CURRENT APPLICATION NUMBER: US/09/444, 053A
; CURRENT FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 55
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-444-053-55

Query Match 0.8%; Score 16; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 AGGCTCAGGAAGCTTG 297
|||||
Db 5 aggcgcaggaagcttg 20

RESULT 16
US-09-457-708-21
; Sequence 21, Application US/09457708
; Patent No. 6326483
; GENERAL INFORMATION:
; APPLICANT: Kwiatkowski, David J.
; APPLICANT: Sampson, Julian R.
; APPLICANT: Povey, Sue
; APPLICANT: van Slegtenhorst, Marjon
; APPLICANT: Halley, Dicky
; TITLE OF INVENTION: Compositions and Methods Based U
; TITLE OF INVENTION: Sclerosis-1 (TSC-1) Gene and Gene Product
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Vinson & Elkins
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20004-1008
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/457, 708
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanzo, Michael A.
; REGISTRATION NUMBER: 36,912
; REFERENCE/DOCKET NUMBER: BRI31/42002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 639-6585
; TELEFAX: (202) 639-6604
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "PCR primer"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-09-457-708-21

Query Match 0.8%; Score 16; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1713 ACCTGCTGAAGAAG 1728
|||||
Db 1 ACCTGCTGAAGAAG 16

RESULT 17
US-08-232-144-2/c
; Sequence 2, Application US/08232144
; Patent No. 5571695
; GENERAL INFORMATION:
; APPLICANT: SELBIE, Lisa
; APPLICANT: HERZOG, Herbert
; APPLICANT: SHINE, John
; TITLE OF INVENTION: Human Neuropeptide Y-Y1 Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Flgg, Ernst & Kurz
; STREET: 555 13th St, N.W., Suite 701-East
; CITY: Washington
; STATE: DC
; COUNTRY: US
; ZIP: 20004

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232, 144
; FILING DATE: 26-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ERNST, Barbara G
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1871-107A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-232-144-2

Query Match 0.8%; Score 16; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1765 CAGAATGGAATTC 1780
|||||
Db 27 CAGAATGGAATTC 12

RESULT 18
US-08-230-002-17
; Sequence 17, Application US/08230002

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,080
FILING DATE: 17-JUL-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/592,126
FILING DATE: 26-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: INTRON 4 OF RAD50 GENOMIC SEQUENCE
US-08-687-080-65

Query Match 0.8%; Score 16; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 612 ATGGACTACAAAGCTT 627
|||||
Db 33 ATGGACTACAAAGCTT 18

RESULT 21
US-08-592-126-137
Sequence 137, Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
TITLE OF INVENTION: Polypeptides
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:
LENGTH: 574 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G115c.seq
US-08-592-126-137

Query Match 0.8%; Score 16; DB 1; Length 574;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1361 TTATGATGAGGAGTG 1376
|||||
Db 342 TTATGATGAGGAGTG 357

RESULT 22
US-08-784-289-1/c
Sequence 1, Application US/08784289
Patent No. 5817912
GENERAL INFORMATION:
APPLICANT: Pedrazzini, Thierry
APPLICANT: Brunner, Hans R.
TITLE OF INVENTION: Transgenic Animals with Disrupted NPY Y1
TITLE OF INVENTION: Receptor Genes
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Vinson & Elkins L.L.P.
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,289
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: BMR 350/13000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 639-6585
TELEFAX: (202) 639-6604
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 595 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-784-289-1

Query Match 0.8%; Score 16; DB 1; Length 595;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1765 CAGAAATGAGAAATTC 1780
|||||
Db 16 CAGAAATGAGAAATTC 1

RESULT 23

US-08-858-207A-116/c
; Sequence 116, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858, 207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 116:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 643 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-858-207A-116

Query Match 0.8%; Score 16; DB 4; Length 643;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 696 AAGAACAGGCTATAA 711
|||||
Db 24 AAGAACAGGCTATAA 9

RESULT 24

US-08-858-207A-31/c
; Sequence 31, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds

NUMBER OF SEQUENCES: 552
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858, 207A
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/017670
FILING DATE: 14-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50475
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-858-207A-31

Query Match 0.8%; Score 16; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 47 GAACGCCGCGCTAATG 62
|||||
Db 694 GAACGCCGCGCTAATG 679

RESULT 25

US-08-936-165A-125
; Sequence 125, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 875 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-936-165A-125.

Query Match 0.8%; Score 16; DB 4; Length 875;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1064 GTTTACTAGTGAGTA 1079
DB 480 GTTATACGCGGATA 495

RESULT 26
US-08-592-126-78
Sequence 78, Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESS:
ADDRESSEE: DeLinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Shultz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 890 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G256.seq
US-08-592-126-78

Query Match 0.8%; Score 16; DB 1; Length 890;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1556 ATCAAGACTCTCAAT 1571
DB 310 ATCAAGACTCTCAAT 325

RESULT 27
US-08-368-236-2
Sequence 2, Application US/08368236
Patent No. 5804408
GENERAL INFORMATION:
APPLICANT: Hagihara, et al.
TITLE OF INVENTION: A METHOD FOR EXPRESSING POLYPEPTIDES
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sherman and Shalloway
STREET: 413 N. Washington Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM
OPERATING SYSTEM: DOS 5.1
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/368,236
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/941,139
FILING DATE: NO. 5804408ember 13, 1992
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Richard Steinberg.
REGISTRATION NUMBER: 26,588
REFERENCE/DOCKET NUMBER: S-2336
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 549-2282
TELEFAX: (703) 836-0106
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1186
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: NO
US-08-368-236-2

Query Match 0.8%; Score 16; DB 1; Length 1186;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 CGTTCTATCGAAGAC 118
DB 708 CGTTCTATCGAAGAC 723

RESULT 28
US-08-245-295-6/C
; Sequence 6, Application US/08245295
; Patent No. 5700658
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Kilgannon, Patrick D.
; TITLE OF INVENTION: ICAM-4 Materials and Methods
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,295
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/32055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1295 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-245-295-6

Query Match 0.8%; Score 16; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTC 206
|||||
DB 723 AATTCCTCCAGGTC 708

RESULT 29
US-08-481-130-6/C
; Sequence 6, Application US/08481130
; Patent No. 5702917
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Kilgannon, Patrick D.
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,130
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32713
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-481-130-6

Query Match 0.8%; Score 16; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTC 206
|||||
DB 723 AATTCCTCCAGGTC 708

RESULT 30
US-08-656-984A-6/C
; Sequence 6, Application US/08656984A
; Patent No. 5753502
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Kilgannon, Patrick D.
; TITLE OF INVENTION: ICAM-4 Materials and Methods

```

; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,984A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,295
; FILING DATE: 18-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,604
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, JR. JOSEPH A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/33321
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1295 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-656-984A-6

Query Match 0.8%; Score 16; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
Db 723 AATTCCTCCAGGTG 708

RESULT 31
; US-08-485-604-6/c
; Sequence 6, Application US/08485604
; Patent No. 5773293
; GENERAL INFORMATION:
; APPLICANT: W.P. W. Michael
; APPLICANT: Kilgannon, Patrick D.
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
```

```

; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,604
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,295
; FILING DATE: 18-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, JR. JOSEPH A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32715
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1295 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-485-604-6

Query Match 0.8%; Score 16; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
Db 723 AATTCCTCCAGGTG 708

RESULT 32
; US-08-487-595-6/c
; Sequence 6, Application US/08487595
; Patent No. 5852170
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Kilgannon, Patrick D.
; TITLE OF INVENTION: ICAM-4 Materials and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
```

STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,595
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32714
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-487-595-6

Query Match 0.8%; Score 16; DB 2; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 191 AATTCCTCCTCAAGTG 206
DB 723 AATTCCTCCTCAAGTG 708
RESULT 33
US-08-863-790-25/c
Sequence 25, Application US/08863790
Patent No. 6087130
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Vazeux, Rosemay
TITLE OF INVENTION: ICAM-Related Protein
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSER: Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago

STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/863,790
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 6087130and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31570
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)474-6300
TELEFAX: (312)474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-863-790-25

Query Match 0.8%; Score 16; DB 3; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 191 AATTCCTCCTCAAGTG 206
DB 723 AATTCCTCCTCAAGTG 708
RESULT 34
US-08-296-749-25/c
Sequence 25, Application US/08296749
Patent No. 6153395
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Vazeux, Rosemay
TITLE OF INVENTION: ICAM-Related Protein
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSER: Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,749
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 615395and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31570
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)474-6300
TELEFAX: (312)474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1295 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-296-749-25

Query Match 0.8%; Score 16; DB 3; Length 1295;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 191 AATTCCTCCAGGTG 206
|||||
DB 723 AATTCCTCCAGGTG 708

RESULT 35
US-08-440-845D-9
Sequence 9, Application US/08440845D
Patent No. 5955329
GENERAL INFORMATION:
APPLICANT: Yuan, L.
APPLICANT: Kitidi, J.
APPLICANT: Dehesh, K.
APPLICANT: Knauf, V.
TITLE OF INVENTION: Engineering Plant Thioesterases for
TITLE OF INVENTION: Altered Substrate Specificity.
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.40 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,845D
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Donna E. Scherer

REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 113
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-440-845D-9

Query Match 0.8%; Score 16; DB 2; Length 1314;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1364 AGTGATGGAGTGAAA 1379
|||||
DB 1058 AGTGATGGAGTGAAA 1073

RESULT 36
US-08-868-458-9
Sequence 9, Application US/08868458
Patent No. 6150512
GENERAL INFORMATION:
APPLICANT: Yuan, L.
TITLE OF INVENTION: Engineering Plant Thioesterases And
TITLE OF INVENTION: Disclosure of Plant Thioesterases
TITLE OF INVENTION: Having No. 6150512el Substrate Specificity
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: IBM PC
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: Microsoft Word for Window 95 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/868,458
FILING DATE: 03-Jun-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/07064
FILING DATE: 15-MAY-96
APPLICATION NUMBER: 08/537,083
FILING DATE: 29-SEPT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 133
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1314 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-08-868-458-9

Query Match 0.8%; Score 16; DB 3; Length 1314;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1364 AGTATGAGAGTGA 1379
1058 AGTATGAGAGTGA 1073

RESULT 37

US-08-611-280-24
; Sequence 24, Application US/08611280
; Patent No. 5891666
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Grossman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSTRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303
; CITY: Mississauga
; STATE: Ontario
; COUNTRY: Canada
; ZIP: L5N 6J8
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,280
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-338A
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1353 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-611-280-24

Query Match 0.8%; Score 16; DB 2; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1727 AGCAATCCAGAGATT 1742
1300 AGCAATCCAGAGATT 1315

RESULT 38

US-09-195-940-24
; Sequence 24, Application US/09195940
; Patent No. 6258935
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Toshifumi
; APPLICANT: Grossman, Alex
; APPLICANT: Richardson, Christopher D.
; TITLE OF INVENTION: NOVEL GENES ENCODING LSTRF POLYPEPTIDES
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Canada Inc.
; STREET: 6733 Mississauga Road, Suite 303

CITY: Mississauga
STATE: Ontario
COUNTRY: Canada
ZIP: L5N 6J8
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/195,940
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/611,280
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-338A
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1353 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-195-940-24

Query Match 0.8%; Score 16; DB 4; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1727 AGCAATCCAGAGATT 1742
1300 AGCAATCCAGAGATT 1315

RESULT 39

US-08-001-078A-2
; Sequence 2, Application US/08001078A
; Patent No. 5872094
; GENERAL INFORMATION:
; APPLICANT: Goetlink, Paul F.
; APPLICANT: Tondravi, M., Mehriad
; APPLICANT: Binette, Francois
; TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/001,078A
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-001-078A-2

Query Match 0.8%; Score 16; DB 2; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 663 ATTACCTCAAGAGAGT 678
|||||
DB 494 ATTACCTCAAGAGAGT 509

RESULT 40

US-08-463-218-2
; Sequence 2, Application US/08463218
; Patent No. 5986052
; GENERAL INFORMATION:
; APPLICANT: Goetlinck, Paul F.
; APPLICANT: Tondravi, Menrad
; APPLICANT: Binette, Francois
; TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, Suite 510
; CITY: BOSTON
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,218
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,078
; FILING DATE: 06-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-008DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-463-218-2

Query Match 0.8%; Score 16; DB 2; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 663 ATTACCTCAAGAGAGT 678
|||||
DB 494 ATTACCTCAAGAGAGT 509

RESULT 41

PCT-US94-00253-2
; Sequence 2, Application PC/TUS9400253
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR PROMOTING CARTILAGE MATRIX
; NUMBER OF SEQUENCES: 2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00253
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,078
; FILING DATE: 06-JAN-1993
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US94-00253-2

Query Match 0.8%; Score 16; DB 5; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 663 ATTACCTCAAGAGAGT 678
|||||
DB 494 ATTACCTCAAGAGAGT 509

RESULT 42

US-08-245-295-9/c
; Sequence 9, Application US/08245295
; Patent No. 5700658
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Kilgannon, Patrick D.
; TITLE OF INVENTION: ICAM-4 Materials and Methods
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, Suite 6300
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,295
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:


```

; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Borun, Michael F.
; REGISTRATION NUMBER: 25,447
; REFERENCE/DOCKET NUMBER: 27866/32055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEEX: 25-3856
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1472 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-245-295-9

Query Match          0.8%; Score 16; DB 1; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
    |||||||||||||||
DB 723 AATTCCTCCAGGTG 708

RESULT 43
US-08-481-130-9/c
; Sequence 9, Application US/08481130
; Patent No. 5702917
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Kilgannon, Patrick D.
; TITLE OF INVENTION: ICAM-4 Materials and Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,130
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,295
; FILING DATE: 18-MAY-1994
; ATTORNEY/AGENT INFORMATION:
;
```

```

; NAME: WILLIAMS, JR. JOSEPH A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32713
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEEX: 25-3856
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1472 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-481-130-9

Query Match          0.8%; Score 16; DB 1; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
    |||||||||||||||
DB 723 AATTCCTCCAGGTG 708

RESULT 44
US-08-656-984A-9/c
; Sequence 9, Application US/08656984A
; Patent No. 5753502
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Kilgannon, Patrick D.
; TITLE OF INVENTION: ICAM-4 Materials and Methods
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,984A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,295
; FILING DATE: 18-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,604
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
;
```

NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33321
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1472 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-656-984A-9

Query Match 0.8%; Score 16; DB 1; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
|||||
Db 723 AATTCCTCCAGGTG 708

RESULT 45
US-08-485-604-9/C
Sequence 9, Application US/08485604
Patent No. 5773293
GENERAL INFORMATION:
APPLICANT: W.P. Michael
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,604
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32715

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1472 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-485-604-9

Query Match 0.8%; Score 16; DB 1; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATTCCTCCAGGTG 206
|||||
Db 723 AATTCCTCCAGGTG 708

RESULT 46
US-08-487-595-9/C
Sequence 9, Application US/08487595
Patent No. 5852170
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: ICAM-4 Materials and Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,595
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,295
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, JR. JOSEPH A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32714
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448

TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1472 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-487-595-9

Query Match 0.8%; Score 16; DB 2; Length 1472;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 AATCCCTCCAGGTG 206
|||||
DB 723 AATCCCTCCAGGTG 708

RESULT 47

US-08-443-568B-15/c
Sequence 15, Application US/08443568B
Patent No. 5759807
GENERAL INFORMATION:
APPLICANT: Breese, Tim
APPLICANT: Havenga, Kirk
APPLICANT: Rindersknecht, Ernst
APPLICANT: Vandien, Richard
APPLICANT: Daniel, Yansura
TITLE OF INVENTION: Process for Producing Relaxin
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443,568B
FILING DATE: 22-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/080,354
FILING DATE: 21-JUNE-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Adams, Samuel B.
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 7842-037
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1500 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 438..1235
US-08-443-568B-15

Query Match 0.8%; Score 16; DB 1; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 874 GCTAGAGCAGCCCAT 889
|||||
DB 1474 GCTAGAGCAGCCCAT 1459

RESULT 48

PCT-US94-06997-15/c
Sequence 15, Application PC/TUS9406997
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: 460 Point San Bruno Boulevard
APPLICANT: South San Francisco, California
APPLICANT: United States of America
TITLE OF INVENTION: Process for Producing Relaxin
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06997
FILING DATE: 20-JUN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 7842-025-228
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1500 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 438..1235
PCT-US94-06997-15

Query Match 0.8%; Score 16; DB 5; Length 1500;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 874 GCTAGAGCAGCCCAT 889
|||||
DB 1474 GCTAGAGCAGCCCAT 1459

RESULT 49
US-07-759-568-4
Sequence 4, Application US/07759568
Patent No. 5374506
GENERAL INFORMATION:
APPLICANT: Murphy, Philip M.
TITLE OF INVENTION: Cloning of cDNA Encoding a Functional
TITLE OF INVENTION: Human Interleukin-8 Receptor

NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman, Darby & Cushman
STREET: 1615 L Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/759,568
FILING DATE: 19910913
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26581
REFERENCE/DOCKET NUMBER: WTS/5683/91535/WHH
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 cush
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1510 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-07-759-568-4

Query Match 0.8%; Score 16; DB 1; Length 1510;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 432 ATTTAAAGTTACCTCA 447
|||||
Db 7 ATTTAAAGTTACCTCA 22

RESULT 50
US-08-288-899-1/C
Sequence 1, Application US/08288899
Patent No. 5610036
GENERAL INFORMATION:
APPLICANT: MURA, MASAMI
APPLICANT: ISHIDA, YUTAKA
APPLICANT: OI, HIDEYUKI
APPLICANT: MURAKAMI, YUKIMITSU
TITLE OF INVENTION: MUTANT AOX2 PROMOTER, MICROORGANISM
TITLE OF INVENTION: CARRYING SAME, METHOD OF PREPARATION THEREOF, AND
TITLE OF INVENTION: PRODUCTION OF HETEROLOGOUS PROTEIN USING SUCH
TITLE OF INVENTION: MICROORGANISM
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn Macpeak and Seas
STREET: 2100 Pennsylvania Avenue, NW
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,899

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/858,830
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Biggart, Waddell A.
REGISTRATION NUMBER: 24,861
REFERENCE/DOCKET NUMBER: 028202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: plasmid DNA
FEATURE:
NAME/KEY: promoter
LOCATION: 1
US-08-288-899-1

Query Match 0.8%; Score 16; DB 1; Length 1528;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 811 AGACGATGATTCGCA 826
|||||
Db 931 AGACGATGATTCGCA 916

Search completed: June 13, 2002, 12:36:10
Job time: 10738 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 11:00:47 : Search time 3439.62 Seconds
(without alignments)
12573.529 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTACGACATATAAA.....ATCCTGACGACCAACCA 1999

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 21979536 seqs, 10817449327 residues

Word size : 15

Total number of hits satisfying chosen parameters: 41001

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/2/pna/PCRTUS_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US080_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/2/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/2/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/2/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/2/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/2/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/2/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/2/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/2/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/2/pna/US095_COMB.seq:*
20: /cgn2_6/ptodata/2/pna/US095A_COMB.seq:*
21: /cgn2_6/ptodata/2/pna/US095B_COMB.seq:*
22: /cgn2_6/ptodata/2/pna/US095C_COMB.seq:*
23: /cgn2_6/ptodata/2/pna/US095D_COMB.seq:*
24: /cgn2_6/ptodata/2/pna/US096_COMB.seq:*
25: /cgn2_6/ptodata/2/pna/US096A_COMB.seq:*
26: /cgn2_6/ptodata/2/pna/US096B_COMB.seq:*
27: /cgn2_6/ptodata/2/pna/US096C_COMB.seq:*
28: /cgn2_6/ptodata/2/pna/US096D_COMB.seq:*
29: /cgn2_6/ptodata/2/pna/US097_COMB.seq:*
30: /cgn2_6/ptodata/2/pna/US097A_COMB.seq:*
31: /cgn2_6/ptodata/2/pna/US097B_COMB.seq:*
32: /cgn2_6/ptodata/2/pna/US097C_COMB.seq:*
33: /cgn2_6/ptodata/2/pna/US098_COMB.seq:*
34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq:*
35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq:*
36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq:*
37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq:*
38: /cgn2_6/ptodata/2/pna/US099_COMB.seq:*
39: /cgn2_6/ptodata/2/pna/US099A_COMB.seq:*
40: /cgn2_6/ptodata/2/pna/US099B_COMB.seq:*
41: /cgn2_6/ptodata/2/pna/US099C_COMB.seq:*
42: /cgn2_6/ptodata/2/pna/US099D_COMB.seq:*
43: /cgn2_6/ptodata/2/pna/US099E_COMB.seq:*

44: /cgn2_6/ptodata/2/pna/US6005_COMB.seq:*
45: /cgn2_6/ptodata/2/pna/US6006_COMB.seq:*
46: /cgn2_6/ptodata/2/pna/US6007_COMB.seq:*
47: /cgn2_6/ptodata/2/pna/US6008_COMB.seq:*
48: /cgn2_6/ptodata/2/pna/US6009_COMB.seq:*
49: /cgn2_6/ptodata/2/pna/US6010_COMB.seq:*
50: /cgn2_6/ptodata/2/pna/US6011_COMB.seq:*
51: /cgn2_6/ptodata/2/pna/US6012_COMB.seq:*
52: /cgn2_6/ptodata/2/pna/US6013_COMB.seq:*
53: /cgn2_6/ptodata/2/pna/US6014_COMB.seq:*
54: /cgn2_6/ptodata/2/pna/US6015_COMB.seq:*
55: /cgn2_6/ptodata/2/pna/US6016_COMB.seq:*
56: /cgn2_6/ptodata/2/pna/US6017_COMB.seq:*
57: /cgn2_6/ptodata/2/pna/US6018_COMB.seq:*
58: /cgn2_6/ptodata/2/pna/US6019_COMB.seq:*
59: /cgn2_6/ptodata/2/pna/US6020_COMB.seq:*
60: /cgn2_6/ptodata/2/pna/US6021_COMB.seq:*
61: /cgn2_6/ptodata/2/pna/US6022_COMB.seq:*
62: /cgn2_6/ptodata/2/pna/US6023_COMB.seq:*
63: /cgn2_6/ptodata/2/pna/US6024_COMB.seq:*
64: /cgn2_6/ptodata/2/pna/US6025_COMB.seq:*
65: /cgn2_6/ptodata/2/pna/US6026_COMB.seq:*
66: /cgn2_6/ptodata/2/pna/US6027_COMB.seq:*
67: /cgn2_6/ptodata/2/pna/US6028_COMB.seq:*
68: /cgn2_6/ptodata/2/pna/US6029_COMB.seq:*
69: /cgn2_6/ptodata/2/pna/US6030_COMB.seq:*
70: /cgn2_6/ptodata/2/pna/US6031_COMB.seq:*
71: /cgn2_6/ptodata/2/pna/US6032_COMB.seq:*
72: /cgn2_6/ptodata/2/pna/US6033_COMB.seq:*
73: /cgn2_6/ptodata/2/pna/US6034_COMB.seq:*
74: /cgn2_6/ptodata/2/pna/US6035_COMB.seq:*
75: /cgn2_6/ptodata/2/pna/US6036_COMB.seq:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999	100.0	1999	US-09-536-784-1	Sequence 1, Appl1
2	1999	100.0	1999	US-09-765-271-1	Sequence 1, Appl1
3	1999	100.0	1999	US-09-765-272-1	Sequence 1, Appl1
4	1999	100.0	10383	US-60-029-960-81	Sequence 81, Appl1
5	1999	100.0	10711	US-08-961-527-145	Sequence 145, Appl1
6	1642	82.1	2160	US-09-583-110-1312	Sequence 1312, Ap
7	1642	82.1	2166	US-09-107-433-1102	Sequence 1102, Ap
8	1438	71.9	2160	PCT-US02-03987-9325	Sequence 9325, Ap
9	1438	71.9	2160	US-09-815-242-9325	Sequence 9325, Ap
10	1438	71.9	2160	US-10-072-651-9325	Sequence 9325, Ap
11	1438	71.9	10333	US-60-061-998-996	Sequence 996, Appl1
12	796	39.8	960	US-08-116-541-3	Sequence 3, Appl1
13	541	27.1	2172	PCT-US97-14436-139	Sequence 139, Appl1
14	541	27.1	2172	US-08-911-503-139	Sequence 139, Appl1
15	541	27.1	2172	US-08-911-503-139	Sequence 139, Appl1
16	293	14.7	9845	US-60-029-960-934	Sequence 934, Appl1
17	223	11.2	323	US-60-029-960-934	Sequence 934, Appl1
18	204	10.2	204	US-09-583-110-1313	Sequence 1313, Ap
19	25	1.3	8395	US-09-134-000-1534	Sequence 1534, Ap
20	25	1.3	8395	US-09-070-927-217	Sequence 217, Appl1
21	25	1.3	8395	US-09-070-927-217	Sequence 217, Appl1
22	23	1.2	33	US-09-536-784-227	Sequence 227, Appl1
23	23	1.2	33	US-09-765-271-227	Sequence 227, Appl1
24	23	1.2	33	US-09-765-272-227	Sequence 227, Appl1
25	22	1.1	933	US-60-045-649-108	Sequence 108, Appl1
26	22	1.1	933	US-60-046-653-183	Sequence 183, Appl1
27	22	1.1	2337	PCT-US02-03987-6631	Sequence 6631, Ap
28	22	1.1	2337	US-09-815-242-6631	Sequence 6631, Ap
29	22	1.1	2337	US-10-072-851-6631	Sequence 6631, Ap
30	22	1.1	3478	US-60-068-217-710	Sequence 710, Appl1
31	21	1.1	537	US-09-649-163-2312	Sequence 2312, Ap

908	18	0.9	32768	62	US-60-234-446-211	Sequence 211, App
c 909	18	0.9	33330	31	US-09-803-736-1303	Sequence 1303, Ap
c 910	18	0.9	36342	32	US-09-855-768-16	Sequence 16, Appl
c 911	18	0.9	37866	35	US-09-948-941-629	Sequence 629, App
c 912	18	0.9	39353	1	PCT-US01-01354-22008	Sequence 22008, A
c 913	18	0.9	39353	1	PCT-US01-01354-22044	Sequence 22044, A
c 914	18	0.9	39353	30	US-09-764-905-20208	Sequence 20208, A
c 915	18	0.9	39353	30	US-09-764-905-22045	Sequence 22045, A
c 916	18	0.9	39353	37	US-10-092-399-22008	Sequence 22008, A
c 917	18	0.9	39353	37	US-10-092-399-22044	Sequence 22044, A
c 918	18	0.9	39358	1	PCT-US01-01354-22009	Sequence 22009, A
c 919	18	0.9	39358	1	PCT-US01-01354-22045	Sequence 22045, A
c 920	18	0.9	39358	30	US-09-764-905-20209	Sequence 20209, A
c 921	18	0.9	39358	30	US-09-764-905-22045	Sequence 22045, A
c 922	18	0.9	39358	37	US-10-092-399-22009	Sequence 22009, A
c 923	18	0.9	39358	37	US-10-092-399-22045	Sequence 22045, A
c 924	18	0.9	42004	63	US-60-243-468-614	Sequence 614, App
c 925	18	0.9	43327	19	US-09-528-237A-2263	Sequence 2263, App
c 926	18	0.9	43593	28	US-09-702-134-4600	Sequence 4600, Ap
c 927	18	0.9	43593	31	US-09-815-264-59331	Sequence 59331, A
c 928	18	0.9	44567	37	US-10-004-113-46	Sequence 46, Appl
c 929	18	0.9	46509	55	US-60-161-932-198	Sequence 198, Appl
c 930	18	0.9	49348	19	US-09-528-237A-1927	Sequence 1927, Ap
c 931	18	0.9	49989	25	US-09-647-377-10	Sequence 10, Appl
c 932	18	0.9	63535	35	US-09-948-933-530	Sequence 530, Appl
c 933	18	0.9	65758	63	US-60-245-227-17	Sequence 17, Appl
c 934	18	0.9	65758	64	US-60-258-278-5	Sequence 5, Appl
c 935	18	0.9	69556	24	US-09-620-332-5588	Sequence 2788, A
c 936	18	0.9	69556	28	US-09-702-134-27788	Sequence 27788, A
c 937	18	0.9	69556	31	US-09-815-264-61364	Sequence 61364, A
c 938	18	0.9	78193	63	US-60-245-224-54	Sequence 54, Appl
c 939	18	0.9	81355	63	US-60-242-679-594	Sequence 594, Appl
c 940	18	0.9	82912	27	US-09-692-412-86	Sequence 86, Appl
c 941	18	0.9	82912	31	US-09-803-736-1066	Sequence 1066, Ap
c 942	18	0.9	84196	27	US-09-692-412-94	Sequence 94, Appl
c 943	18	0.9	84196	31	US-09-803-736-1074	Sequence 1074, Ap
c 944	18	0.9	84974	27	US-09-692-412-7	Sequence 7, Appl
c 945	18	0.9	84974	31	US-09-803-736-987	Sequence 987, Appl
c 946	18	0.9	88784	55	US-60-161-932-397	Sequence 397, App
c 947	18	0.9	96501	19	US-09-528-237A-1333	Sequence 1333, App
c 948	18	0.9	97700	64	US-60-257-537-210	Sequence 210, Appl
c 949	18	0.9	105644	27	US-09-692-412-92	Sequence 92, Appl
c 950	18	0.9	105644	31	US-09-803-736-1072	Sequence 1072, Ap
c 951	18	0.9	109786	20	US-09-534-853-543	Sequence 543, App
c 952	18	0.9	109786	31	US-09-803-736-543	Sequence 543, App
c 953	18	0.9	114759	27	US-09-692-412-118	Sequence 118, App
c 954	18	0.9	114759	31	US-09-803-736-1098	Sequence 1098, Ap
c 955	18	0.9	119430	20	US-09-534-859-617	Sequence 617, App
c 956	18	0.9	119430	31	US-09-803-736-617	Sequence 617, App
c 957	18	0.9	124209	62	US-60-236-804-64	Sequence 64, Appl
c 958	18	0.9	127968	31	US-09-803-736-1127	Sequence 1127, Ap
c 959	18	0.9	131692	27	US-09-692-412-21	Sequence 21, Appl
c 960	18	0.9	131692	31	US-09-803-736-1001	Sequence 1001, Ap
c 961	18	0.9	136328	61	US-60-226-176-2256	Sequence 2256, Ap
c 962	18	0.9	136328	62	US-60-233-468-2256	Sequence 2256, Ap
c 963	18	0.9	136328	70	US-60-313-371-2256	Sequence 2256, Ap
c 964	18	0.9	149310	64	US-60-212-164-133	Sequence 133, Appl
c 965	18	0.9	158434	60	US-60-254-168-36	Sequence 168, App
c 966	18	0.9	165951	19	US-09-528-237A-1761	Sequence 1761, Ap
c 967	18	0.9	179651	64	US-60-257-537-345	Sequence 345, App
c 968	18	0.9	240825	30	US-09-790-289-1	Sequence 1, Appl
c 969	18	0.9	259000	19	US-09-528-237A-1938	Sequence 1938, Ap
c 970	18	0.9	277797	19	US-09-528-237A-1224	Sequence 1224, Ap
c 971	18	0.9	312957	35	US-09-948-947-18	Sequence 128, App
c 972	18	0.9	312972	35	US-09-948-947-157	Sequence 157, App
c 973	18	0.9	666448	14	US-09-012-031-312	Sequence 312, App
c 974	18	0.9	666448	14	US-09-012-031B-312	Sequence 312, App
c 975	18	0.9	666448	14	US-09-012-031C-312	Sequence 312, App
c 976	18	0.9	666448	17	US-09-335-032-12214	Sequence 12214, A
c 977	18	0.9	713059	24	US-09-634-306B-174561	Sequence 174561, A
c 978	18	0.9	924430	14	US-09-012-031B-314	Sequence 314, App
c 979	18	0.9	924430	14	US-09-012-031C-314	Sequence 314, App
c 980	18	0.9	924430	14	US-09-012-031C-314	Sequence 314, App

c 981	18	0.9	924430	17	US-09-335-032-12216	Sequence 12216, A
c 982	17	0.9	25	26	US-09-660-220-7345	Sequence 7345, Ap
c 983	17	0.9	25	26	US-09-956-584-381942	Sequence 381942, A
c 984	17	0.9	25	62	US-60-234-017-356007	Sequence 356007, A
c 985	17	0.9	25	74	US-60-353-987-596126	Sequence 596126, A
c 986	17	0.9	25	74	US-60-353-987-786733	Sequence 786733, A
c 987	17	0.9	27	19	US-09-526-597A-21	Sequence 21, Appl
c 988	17	0.9	27	19	US-09-526-597A-14	Sequence 14, Appl
c 989	17	0.9	27	20	US-09-574-735C-14	Sequence 228, App
c 990	17	0.9	27	22	US-09-574-735C-14	Sequence 14, Appl
c 991	17	0.9	27	22	US-09-574-735C-20	Sequence 20, Appl
c 992	17	0.9	27	30	US-09-765-271-228	Sequence 228, App
c 993	17	0.9	27	30	US-09-765-272-1228	Sequence 228, App
c 994	17	0.9	35	17	US-09-385-222A-7	Sequence 7, Appl
c 995	17	0.9	35	17	US-09-385-222A-7	Sequence 7, Appl
c 996	17	0.9	36	17	US-09-385-222A-6	Sequence 6, Appl
c 997	17	0.9	36	17	US-09-385-222A-6	Sequence 6, Appl
c 998	17	0.9	51	56	US-60-172-377-6188	Sequence 6188, Ap
c 999	17	0.9	51	56	US-60-172-377-7479	Sequence 7479, Ap
c 1000	17	0.9	112	14	US-09-076-667-3248	Sequence 3248, Ap

ALIGNMENTS

```

RESULT 1
US-09-536-784-1
; Sequence 1, Application US/09536784
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536, 784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-536-784-1

Query Match 100.0%; Score 1999; DB 20; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TAAATCTACGACATAAATCACTTCATTCGCTGAGCTTGGCTTCGACGCGCGCTCA 60

```

|||||
Db 1 TAAATCTAGCAATAAAAAATCACTATGCTGACTGGGTTTGAAAGCGCGCTCA 60
Qy 61 TGGCCAGGTAAATGATTTCCACAGATTTGGTTAAGGAATCGTTTAAATGAGAACCA 120
Db 61 TGGCCAGGTAAATGATTTCCACAGATTTGGTTAAGGAATCGTTTAAATGAGAACCA 120
Qy 121 TGGCTTCTGACACAGAGGGGATTTGATACATCCGTATCTGGAGCTTTCTTGGCAA 180
Db 121 TGGCTTCTGACACAGAGGGGATTTGATACATCCGTATCTGGAGCTTTCTTGGCAA 180
Qy 181 TGTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCAACAGTTGATTAAGTTGAC 240
Db 181 TGTGCAAGCAATTCCTCCAAAGTGGATCAACTCTCACCAACAGTTGATTAAGTTGAC 240
Qy 241 TTAAGTTTCACTGAGCTTGGACAGACTATTTCTGTAAGGCTCAGAGAGCTTGAT 300
Db 241 TTAAGTTTCACTGAGCTTGGACAGACTATTTCTGTAAGGCTCAGAGAGCTTGAT 300
Qy 301 ACCGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTACCTATATATAATA 360
Db 301 ACCGATTCAGTTAGAACAAAGCAACCAAGCAAGAAATCTTACCTATATATAATA 360
Qy 361 GGTCTACATGCTTAATGGGACTATGGAATGACAGACAGAGCTCAAACTACTATGATA 420
Db 361 GGTCTACATGCTTAATGGGACTATGGAATGACAGACAGAGCTCAAACTACTATGATA 420
Qy 421 AACACCATATATTAAGTTTACCTACAGTTAGCTTGGCTGGAGATCCCTACAGCAC 480
Db 421 AACACCATATATTAAGTTTACCTACAGTTAGCTTGGCTGGAGATCCCTACAGCAC 480
Qy 481 AACCAATATGACCCCTATTCACATCCAGACAGAGCCAGAGCCGCAAGCTTGCTT 540
Db 481 AACCAATATGACCCCTATTCACATCCAGACAGAGCCAGAGCCGCAAGCTTGCTT 540
Qy 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAAGCACTATAC 600
Db 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAAGCACTATAC 600
Qy 601 ACCAATATGATGAGCTCAAAAGTCTCAAAATCAGCAATATTCCTGCTTACATGGA 660
Db 601 ACCAATATGATGAGCTCAAAAGTCTCAAAATCAGCAATATTCCTGCTTACATGGA 660
Qy 661 TAATTAACCTCAAGAGATCATCAATCAATGTAAGAAAGAAAGAGCTATACCTACTCAC 720
Db 661 TAATTAACCTCAAGAGATCATCAATCAATGTAAGAAAGAAAGAGCTATACCTACTCAC 720
Qy 721 AACTGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
Db 721 AACTGGATGATGCTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
Qy 781 TTAACAATNAGAGCAATAGTGGCTATCCAGACAGATGAATGCAAGTGGCTTACCAAT 840
Db 781 TTAACAATNAGAGCAATAGTGGCTATCCAGACAGATGAATGCAAGTGGCTTACCAAT 840
Qy 841 TGTGATGTTTAAAGGTAAGTCAATGCCAGTAGAGAGCAGCATCAAGTAA 900
Db 841 TGTGATGTTTAAAGGTAAGTCAATGCCAGTAGAGAGCAGCATCAAGTAA 900
Qy 901 TGTTCCTTCGGAATTAACCAAGAGTAGAAAACCCGAGCTGGGATCAACTATGAA 960
Db 901 TGTTCCTTCGGAATTAACCAAGAGTAGAAAACCCGAGCTGGGATCAACTATGAA 960
Qy 961 ACCGATTCAGACTATGCTCCCTGCTGAGTAGAGTGTCTACAGATTCACCTACTAT 1020
Db 961 ACCGATTCAGACTATGCTCCCTGCTGAGTAGAGTGTCTACAGATTCACCTACTAT 1020
Qy 1021 CGTTACAGATGAGCCCTATATACCTGGAGCAATATCTCGTTTAAATGAGGATAG 1080
Db 1021 CGTTACAGATGAGCCCTATATACCTGGAGCAATATCTCGTTTAAATGAGGATAG 1080
Qy 1081 GGGCTACTTTGGCAACATCAGCTTGCAATAGCCCTGCAACAATGCGAAGAGTCCAC 1140
Db 1081 GGGCTACTTTGGCAACATCAGCTTGCAATAGCCCTGCAACAATGCGAAGAGTCCAC 1140

Db 1081 GGGCTACTTTGGCAACATCAGCTTGCAATAGCCCTGCAACAATGCGAAGAGTCCAC 1140
Qy 1141 CGTGAACCTTAACCAAGGTGAGCTCAACCGCGCAAGACTTTCTAAATGAGTACG 1200
Db 1141 CGTGAACCTTAACCAAGGTGAGCTCAACCGCGCAAGACTTTCTAAATGAGTACG 1200
Qy 1201 AATGCACTACCCCAAGTATTCAGTACTCAAAATGCCATTTCAAGTAAACCAACGAATCAGA 1260
Db 1201 AATGCACTACCCCAAGTATTCAGTACTCAAAATGCCATTTCAAGTAAACCAACGAATCAGA 1260
Qy 1261 CAAAAAATATGAGCAAGTAGTGAAGAAATGGCTGCTTACAGCTTGGCAATAG 1320
Db 1261 CAAAAAATATGAGCAAGTAGTGAAGAAATGGCTGCTTACAGCTTGGCAATAG 1320
Qy 1321 TGGAACTTACTTAACCAATGATATTCATTAAGTGAAGCTTATGATGGAGTGA 1380
Db 1321 TGGAACTTACTTAACCAATGATATTCATTAAGTGAAGCTTATGATGGAGTGA 1380
Qy 1381 AGAGTTCTTAATGTCGGAAGTGTGCAATGGAAGGAAAGACAGCCTATATGATGACCA 1440
Db 1381 AGAGTTCTTAATGTCGGAAGTGTGCAATGGAAGGAAAGACAGCCTATATGATGACCA 1440
Qy 1441 CATGATGAAGAACAGTCTTACCTATGGAACCTGGACGAATGCTATCTTGGCTCCC 1500
Db 1441 CATGATGAAGAACAGTCTTACCTATGGAACCTGGACGAATGCTATCTTGGCTCCC 1500
Qy 1501 TCAGGCTGTAAACAGGAAGCTTACATATACAGAGGAAATGAAACCAACATCA 1560
Db 1501 TCAGGCTGTAAACAGGAAGCTTACATATACAGAGGAAATGAAACCAACATCA 1560
Qy 1561 GACCTTCTCAATTTGTAGCACTGATGAACTATTTGCTGCTATACGCTAAATATTCAT 1620
Db 1561 GACCTTCTCAATTTGTAGCACTGATGAACTATTTGCTGCTATACGCTAAATATTCAT 1620
Qy 1621 GCGTATGAGACAGGCTATTTACCGCTGACACCACTTTGAGGAATGGCCCTTACGG 1680
Db 1621 GCGTATGAGACAGGCTATTTACCGCTGACACCACTTTGAGGAATGGCCCTTACGG 1680
Qy 1681 CGCTGCCAAGTTTACCGCTATGATGACCTGCTGGAAGGAAGCAATCCAGAA 1740
Db 1681 CGCTGCCAAGTTTACCGCTATGATGACCTGCTGGAAGGAAGCAATCCAGAA 1740
Qy 1741 TTGGAATATACAGAGGGGCTTACAGAAATGAGAAATGGAATTTAAAAATGGTCTCG 1800
Db 1741 TTGGAATATACAGAGGGGCTTACAGAAATGAGAAATGGAATTTAAAAATGGTCTCG 1800
Qy 1801 TTCTACGTGGAAGTCACTGCTGCTCACACACACCCCATCAAGTGAAGTCAAGTCAATC 1860
Db 1801 TTCTACGTGGAAGTCACTGCTGCTCACACACACCCCATCAAGTGAAGTCAAGTCAATC 1860
Qy 1861 ATCAGATATTCACCTTCAAGTGTGCAACCACTCCAGACAAATTAATGTACGAC 1920
Db 1861 ATCAGATATTCACCTTCAAGTGTGCAACCACTCCAGACAAATTAATGTACGAC 1920
Qy 1921 TACCAATCTTAACCAATATAGCAACCAATATACCAACCCCTGATCAACAAATCAGAA 1980
Db 1921 TACCAATCTTAACCAATATAGCAACCAATATACCAACCCCTGATCAACAAATCAGAA 1980
Qy 1981 TCCTCAACGACACACCA 1999
Db 1981 TCCTCAACGACACACCA 1999

RESULT 2
US-09-765-271-1
; Sequence 1, Application us/09765271
; GENERAL INFORMATION:
; APPLICANT: Chai et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue


```

? CITY: Rockville
? STATE: Maryland
? COUNTRY: USA
? ZIP: 20850
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
? COMPUTER: HP Vectra 486/33
? OPERATING SYSTEM: MSDOS version 6.2
? SOFTWARE: ASCII Text
?
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/765,271
? FILING DATE: 22-Jan-2001
? CLASSIFICATION: <Unknown>
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 09/536,784
? FILING DATE: <Unknown>
? APPLICATION NUMBER: 08/961,083
? FILING DATE: OCT-30-1997
?
? ATTORNEY/AGENT INFORMATION:
? NAME: Michelle S. Marks
? REGISTRATION NUMBER: 41,971
? REFERENCE/DOCKET NUMBER: PB340P3
?
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (301) 309-8504
? TELEFAX: (301) 309-8512
?
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1999 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
?
? SEQUENCE DESCRIPTION: SEQ ID NO: 1:
?
? US-09-765-271-1

```

```

Query Match      100.0%; Score 1999; DB 30; Length 1999;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TAAATCTAGACAAATAAATCACTGATGCTGGGTTCTGAAGCCGCTCAA 60
DB 1 TAAATCTAGACAAATAAATCACTGATGCTGGGTTCTGAAGCCGCTCAA 60
QY 61 TGCCCAAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTCTATCGAAGCA 120
DB 61 TGCCCAAGCTAATGATATTCGCCACAGATTTGGTTAAGGCAATCGTTCTATCGAAGCA 120
QY 121 TCGCTTCTTGACACAGGGGGGATTTGATACCATCGTATCGTGGAGCTTCTTGCGCAA 180
DB 121 TCGCTTCTTGACACAGGGGGGATTTGATACCATCGTATCGTGGAGCTTCTTGCGCAA 180
QY 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCACCACAGTTGATTAAGTTGAC 240
DB 181 TCTGCAAGCAATTCCTCCCAAGGTGATCACTCTCACCACAGTTGATTAAGTTGAC 240
QY 241 TTAATCTTCAACTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT 300
DB 241 TTAATCTTCAACTCGACTTCGACACAGACTATTTCTCGTAAGGCTCAGGAAGCTTGTT 300
QY 301 AGCGATTCACTTAGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 360
DB 301 AGCGATTCACTTAGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 360
QY 361 GGTCTACATGCTCTAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 420
DB 361 GGTCTACATGCTCTAATGGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 420
QY 421 AGACCTCAATTAATTAAGTTACCTCAGTAGAGCTTGCTGCTGGAATGCTCAGGACACC 480
DB 421 AGACCTCAATTAATTAAGTTACCTCAGTAGAGCTTGCTGCTGGAATGCTCAGGACACC 480
QY 481 AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGCCGCAAACTTGCTTT 540
DB 481 AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGCCGCAAACTTGCTTT 540

```

```

DB 481 AAACCAATATGACCCCTATTACATCCAGAGCAGCCCAAGCCGCAAACTTGCTTT 540
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAACAGATGAGAAAGCAATGATAC 600
DB 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGAACAGATGAGAAAGCAATGATAC 600
QY 601 ACCAATTACTGATGAGTACCAAAAGCTCTCAATTCAGCAAGTAATTAATTAATTAATTA 660
DB 601 ACCAATTACTGATGAGTACCAAAAGCTCTCAATTCAGCAAGTAATTAATTAATTAATTA 660
QY 661 TAATTACTCAGAGCAATCATCATCAAGTTGAAGAAAGAAAGCAAGTATTAATTAATTA 720
DB 661 TAATTACTCAGAGCAATCATCATCAAGTTGAAGAAAGAAAGCAAGTATTAATTAATTA 720
QY 721 AACTGGATGATGTCTACACAAATGTAAGCAAGAAAGCTCAAAATATCTGTGGATAT 780
DB 721 AACTGGATGATGTCTACACAAATGTAAGCAAGAAAGCTCAAAATATCTGTGGATAT 780
QY 781 TTACAATACAGACGAATACGTTGCCATCCAGAGATGAATTTGCAAGTCTGCTTACCAT 840
DB 781 TTACAATACAGACGAATACGTTGCCATCCAGAGATGAATTTGCAAGTCTGCTTACCAT 840
QY 841 TGTGATGTTTCTAACGCTAATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 TGTGATGTTTCTAACGCTAATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 TGTTCCTGGAATTAACCAAGCAGTAAGAAACCAACCGGAGTGGGATCAATTAATGA 960
DB 901 TGTTCCTGGAATTAACCAAGCAGTAAGAAACCAACCGGAGTGGGATCAATTAATGA 960
QY 961 ACCGATCAGACATATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 ACCGATCAGACATATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 CGTTCAGATGAGCCCTATTAATACCTCGGACAAATATCTGTTAATTAATTAATTA 1080
DB 1021 CGTTCAGATGAGCCCTATTAATACCTCGGACAAATATCTGTTAATTAATTAATTA 1080
QY 1081 GGGCTACTTTGGCAACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 GGGCTACTTTGGCAACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 CGTGGAACTCTTAACCAAGGCTGAGCTCAACCGGCAAGCTTCTTAATTAATTAATTA 1200
DB 1141 CGTGGAACTCTTAACCAAGGCTGAGCTCAACCGGCAAGCTTCTTAATTAATTAATTA 1200
QY 1201 AATGACATACCAAGTATTAATCACTCAATTAATTAATTAATTAATTAATTAATTA 1260
DB 1201 AATGACATACCAAGTATTAATCACTCAATTAATTAATTAATTAATTAATTAATTA 1260
QY 1261 CAAAAAATATGAGCAAGTATGAAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 1261 CAAAAAATATGAGCAAGTATGAAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1321 TGAACCTTACTATAACCAATGATATTCATTAAGTGTCTTAAAGTGGAGTGA 1380
DB 1321 TGAACCTTACTATAACCAATGATATTCATTAAGTGTCTTAAAGTGGAGTGA 1380
QY 1381 AGAATCTCTAATGCGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
DB 1381 AGAATCTCTAATGCGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1441 CATGATGAAAAACATCTTGAATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500
DB 1441 CATGATGAAAAACATCTTGAATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500
QY 1501 TCAGGCTGTTAAACAGAACTCTAATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGA 1560
DB 1501 TCAGGCTGTTAAACAGAACTCTAATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGA 1560
QY 1561 GACCTCTCAATTTGTACACCTGATGAATTAATTAATTAATTAATTAATTAATTAATTA 1620
DB 1561 GACCTCTCAATTTGTACACCTGATGAATTAATTAATTAATTAATTAATTAATTAATTA 1620

```


QY 1621 GCCTGATGACAGGCTATTCTAACGCTGTGACACCACTTGTAGGCAATGCCCTTACGGT 1680
 DB 1621 GCCTGATGACAGGCTATTCTAACGCTGTGACACCACTTGTAGGCAATGCCCTTACGGT 1680
 QY 1681 CGCTGCCAAGTTTACCGGCTCTATGATGACCTACCTGTCTGGAAGAACATCCAGAGA 1740
 DB 1681 CGCTGCCAAGTTTACCGGCTCTATGATGACCTACCTGTCTGGAAGAACATCCAGAGA 1740
 QY 1741 TTGGATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTCTCG 1800
 DB 1741 TTGGATATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTTAAAAATGGTCTCG 1800
 QY 1801 TTCTAGTGAACCTACCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 1860
 DB 1801 TTCTAGTGAACCTACCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTCATC 1860
 QY 1861 ATCAGATAGTCACTCACTGCTAGCTCAACCACTCCAGACAAATATATAGTAGCAG 1920
 DB 1861 ATCAGATAGTCACTCACTGCTAGCTCAACCACTCCAGACAAATATATAGTAGCAG 1920
 QY 1921 TACCAATCTTAACATATATAGCAATCAATATACACCCCTGATCAACAAATCAGAA 1980
 DB 1921 TACCAATCTTAACATATATAGCAATCAATATACACCCCTGATCAACAAATCAGAA 1980
 QY 1981 TCCTCAACGACGACCA 1999
 DB 1981 TCCTCAACGACGACCA 1999

RESULT 3 US-09-765-272-1

; Sequence 1, Application US/09765272
 ; GENERAL INFORMATION:
 ; APPLICANT: Choi et. al.
 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
 ; NUMBER OF SEQUENCES: 452
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/765,272
 ; FILING DATE: 22-Jan-2001
 ; CLASSIFICATION: <unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/961,083
 ; FILING DATE: <unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1999 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-765-272-1

Query Match 100.0%; Score 1999; DB 30; Length 1999;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTACGACATTAATAATCACTATGCTGACTGGGTTCTGAAACGCCGCTCAA 60
 DB 1 TAAATCTACGACATTAATAATCACTATGCTGACTGGGTTCTGAAACGCCGCTCAA 60
 QY 61 TGGCCAGCAATATGATATCCACAGATTTGGTTAAGGCAATCGTTTATGCAACCA 120
 DB 61 TGGCCAGCAATATGATATCCACAGATTTGGTTAAGGCAATCGTTTATGCAACCA 120
 QY 121 TGGCTTCTTCGACACAGAGGGGATGATACACCTATCCCTGAGAGCTTTCTTGCGCA 180
 DB 121 TGGCTTCTTCGACACAGAGGGGATGATACACCTATCCCTGAGAGCTTTCTTGCGCA 180
 QY 181 TCTGCAAGCAATTCCTCCAGAGTGATCACTTCACCCACAGTGTATTAAGTTGAC 240
 DB 181 TCTGCAAGCAATTCCTCCAGAGTGATCACTTCACCCACAGTGTATTAAGTTGAC 240
 QY 241 TTACTTTTCACTTCGACCTTCGACCACTATTTCTCGTAAGGCTCAGAACCTGGTT 300
 DB 241 TTACTTTTCACTTCGACCTTCGACCACTATTTCTCGTAAGGCTCAGAACCTGGTT 300
 QY 301 AGCGATTCACTTAGAACAACCAACCAAGCAAGAAATCTTACCTACTATTAATAA 360
 DB 301 AGCGATTCACTTAGAACAACCAACCAAGCAAGAAATCTTACCTACTATTAATAA 360
 QY 361 GGTCTACATGTCTAATGGGAATATGGAATGAGAGAGCTCAAACTACTATGTTAA 420
 DB 361 GGTCTACATGTCTAATGGGAATATGGAATGAGAGAGCTCAAACTACTATGTTAA 420
 QY 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGCTGCTGGAATGCTCAGCAC 480
 DB 421 AGACCTCAATTAATTTAAGTTTACCTCAGTTAGCCTTGCTGCTGGAATGCTCAGCAC 480
 QY 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCGCGAAACTTGGTCTT 540
 DB 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCGCGAAACTTGGTCTT 540
 QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC 600
 DB 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTCAATAC 600
 QY 601 ACCAATTAATGATGACCTCAAAAGTCTCAATTCAGCAAGTAATTAACCTCCTTACATGGA 660
 DB 601 ACCAATTAATGATGACCTCAAAAGTCTCAATTCAGCAAGTAATTAACCTCCTTACATGGA 660
 QY 661 TAATTACTCAAGAGATCATCATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTACAC 720
 DB 661 TAATTACTCAAGAGATCATCATCAAGTTGAAGAAAGCAAGGCTATTAACCTACTACAC 720
 QY 721 AACTGGATGATGCTCTACACAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
 DB 721 AACTGGATGATGCTCTACACAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780
 QY 781 TTACATATACAGAGATTAAGTGGCTATCCAGACATGAATGCAAGTGGCTTCTACAT 840
 DB 781 TTACATATACAGAGATTAAGTGGCTATCCAGACATGAATGCAAGTGGCTTCTACAT 840
 QY 841 TGTGATGTTTCTAACGGTAAGTCTATGCCCCAGCTAGAGACAGCATCAGTCAAGTAA 900
 DB 841 TGTGATGTTTCTAACGGTAAGTCTATGCCCCAGCTAGAGACAGCATCAGTCAAGTAA 900
 QY 901 TGTTCCTTCGGAATTAACCAAGCACTAGAAACAAACCCGAGCTGGGATCAACTATGAA 960
 DB 901 TGTTCCTTCGGAATTAACCAAGCACTAGAAACAAACCCGAGCTGGGATCAACTATGAA 960
 QY 961 ACCGATCAGAGACTATGCTTCGCTTGAAGTACGCTGCTACAGTTCACATGCTACTAT 1020
 DB 961 ACCGATCAGAGACTATGCTTCGCTTGAAGTACGCTGCTACAGTTCACATGCTACTAT 1020
 QY 1021 CGTTCAGATGAGCCCTATTAACCTGGGACAAATTAATCTCCGTGTTAATGAGATG 1080

Dd	1021	CGTTCAGATGAGSCCTATATACCTACCGTGGGAGCAAAATACCTCCTTTATATACCTGGGATG	1080
Qy	1081	GGGCTACTTTGGCACATCATCCTTGGCAATTCGGCCTGGCAACAATTCGGAAAGCTCCAGC	1140
Dd	1081	GGGCTACTTTGGCACATCATCCTTGGCAATTCGGCCTGGCAACAATTCGGAAAGCTCCAGC	1140
Qy	1141	CGTGGAAACTCTAAACAGAGTGGAGTCGAACCGCGCAGACACTTCTCTAAATGGTCTAGG	1200
Dd	1141	CGTGGAAACTCTAAACAGAGTGGAGTCGAACCGCGCAGACACTTCTCTAAATGGTCTAGG	1200
Qy	1201	AATGCATCCCAAGTATTCATCTACTCAAAATGGCAATTTCAAGTACAGACAACCGAATTCGA	1260
Dd	1201	AATGCATCCCAAGTATTCATCTACTCAAAATGGCAATTTCAAGTACAGACAACCGAATTCGA	1260
Qy	1261	CAAAAAATATGAGCAAGTAGTGAAGAAAGTGGCTGGCTAGCGTCCCTTTCGAAATGG	1320
Dd	1261	CAAAAAATATGAGCAAGTAGTGAAGAAAGTGGCTGGCTAGCGTCCCTTTCGAAATGG	1320
Qy	1321	TGGAACTTACTATTAACCAATGATATATTCATATAAGTCGTTTATAGTATGGAGTGAATA	1380
Dd	1321	TGGAACTTACTATTAACCAATGATATATTCATATAAGTCGTTTATAGTATGGAGTGAATA	1380
Qy	1381	AAGATTCTCTAATATGTCGGAACCTGGTCCCATGAAGGAAAGCAGACGCTATATGATGACCGA	1440
Dd	1381	AAGATTCTCTAATATGTCGGAACCTGGTCCCATGAAGGAAAGCAGACGCTATATGATGACCGA	1440
Qy	1441	CATGATGAAAGAACGTCCTATCTATGGAAGTGGAGCAACAAATGCTATCTTGGTGGCTCC	1500
Dd	1441	CATGATGAAAGAACGTCCTATCTATGGAAGTGGAGCAACAAATGCTATCTTGGTGGCTCC	1500
Qy	1501	TCAGGCTGGTAAACACAGGAACCTCTAATCTATACAGACAGGAAATTTGAAACCAACATCAA	1560
Dd	1501	TCAGGCTGGTAAACACAGGAACCTCTAATCTATACAGACAGGAAATTTGAAACCAACATCAA	1560
Qy	1561	GACCTCTCAATTTGTAGACGCTGATAGACATATTTGGCGCTATACGGGTAAATATTCAT	1620
Dd	1561	GACCTCTCAATTTGTAGACGCTGATAGACATATTTGGCGCTATACGGGTAAATATTCAT	1620
Qy	1621	GGCTGATGAGACAGGCTATTTCTAACCGCTGACACCACTTGTAGGCAATGGCCTTACGCT	1680
Dd	1621	GGCTGATGAGACAGGCTATTTCTAACCGCTGACACCACTTGTAGGCAATGGCCTTACGCT	1680
Qy	1681	CGCTGGCAAGTTTACCCTCTATGATGACCTAAGCTGTGGAAGGAACATCCAGAGA	1740
Dd	1681	CGCTGGCAAGTTTACCCTCTATGATGACCTAAGCTGTGGAAGGAACATCCAGAGA	1740
Qy	1741	TTGGAATATACAGAGAGGCTCTACAGAAATGAGAAATTCGTTTAAAAATGGTGGCTCG	1800
Dd	1741	TTGGAATATACAGAGAGGCTCTACAGAAATGAGAAATTCGTTTAAAAATGGTGGCTCG	1800
Qy	1801	TTTCTACGTGGAACTACCTGCTCTCACACACACCCCATCATCACTGAAAGTTCAAGCTCATC	1860
Dd	1801	TTTCTACGTGGAACTACCTGCTCTCACACACACCCCATCATCACTGAAAGTTCAAGCTCATC	1860
Qy	1861	ATCAGATAGTTCAACTTCACAGCTGAGCTGACACACCTCCAGAGCAAAATATATGTACGAC	1920
Dd	1861	ATCAGATAGTTCAACTTCACAGCTGAGCTGACACACCTCCAGAGCAAAATATATGTACGAC	1920
Qy	1921	TACCAATCTTAAACAATATATAGCAACAATCAATATACACCCCTGATCAACAATAATGAGAA	1980
Dd	1921	TACCAATCTTAAACAATATATAGCAACAATCAATATACACCCCTGATCAACAATAATGAGAA	1980
Qy	1981	TTCTCAACAGACACAACCA 1999	
Dd	1981	TTCTCAACAGACACAACCA 1999	

```

OY 481 AAACAAATGAGCCCTATTCATCCAGAGACGCCCAAGCCGGAAGCTGGCTT 540
    |||
DB 9283 AAACCAATATGAGCCCTATTCATCCAGAGAGAGCCCAAGCCGGAAGCTGGCTT 9224
OY 541 ATCTGAAATGAAAAATCAAGGCTACTCTCTGTGACAGTATGAGAAAAGCACTCATAC 600
    |||
DB 9223 ATCTGAAATGAAAAATCAAGGCTACTCTCTGTGACAGTATGAGAAAAGCACTCATAC 9164
OY 601 ACCAATATGATGAGCTCAAGAGTCTCAATCAGCAATATTAATCCGTTACTATGGA 660
    |||
DB 9163 ACCAATATGATGAGCTCAAGAGTCTCAATCAGCAATATTAATCCGTTACTATGGA 9104
OY 661 TAATTAACCTCAAGAGATCATCATCAAGTTGAGAGAGAGAGAGCTATTAACCTACTCAC 720
    |||
DB 9103 TAATTAACCTCAAGAGATCATCATCAAGTTGAGAGAGAGAGAGCTATTAACCTACTCAC 9044
OY 721 AACTGGGATGATGCTCTACACAAATGTAGACAGAGAGCTCAAAAACATCTGTGGATAT 780
    |||
DB 9043 AACTGGGATGATGCTCTACACAAATGTAGACAGAGAGCTCAAAAACATCTGTGGATAT 8984
OY 781 TTACATATACAGAGATAGCTGCTGCTATCCAGAGAGAGAGAGAGAGCTGCTTACCAT 840
    |||
DB 8983 TTACATATACAGAGATAGCTGCTGCTATCCAGAGAGAGAGAGAGAGCTGCTTACCAT 8924
OY 841 TGTGATGCTTCTCAAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
    |||
DB 8923 TGTGATGCTTCTCAAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8864
OY 901 TGTGATGCTTCTCAAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
    |||
DB 8863 TGTGATGCTTCTCAAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8804
OY 961 ACCGATACAGAGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
    |||
DB 8803 ACCGATACAGAGATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8744
OY 1021 CGTTACAGATGAGCCCTATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
    |||
DB 8743 CGTTACAGATGAGCCCTATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8684
OY 1081 GGGCTACTTGGCAACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
    |||
DB 8683 GGGCTACTTGGCAACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8624
OY 1141 CGTTGAACTTAACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1200
    |||
DB 8623 CGTTGAACTTAACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8564
OY 1201 AATGCACTACCAAGATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
    |||
DB 8563 AATGCACTACCAAGATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8504
OY 1261 CAAAAAATATGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320
    |||
DB 8503 CAAAAAATATGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8444
OY 1321 TGGAACTTACTTAACCAATGATATATCAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
    |||
DB 8443 TGGAACTTACTTAACCAATGATATATCAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8384
OY 1381 AGAGTTCTCTAATGCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1440
    |||
DB 8383 AGAGTTCTCTAATGCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8324
OY 1441 CATGATGAAAGAGTCTGATTAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1500
    |||
DB 8323 CATGATGAAAGAGTCTGATTAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8264
OY 1501 TCAGGCTGTTAAACAGAGAGTCTGATTAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1560
    |||
DB 8263 TCAGGCTGTTAAACAGAGAGTCTGATTAATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8204
OY 1561 GACCTCTCAATTTGAGACCTGATGAACATATTTGCTGGCTATAGCGCTTAATATTTCAAT 1620

```

```

DB 8203 GACCTCTCAATTTGAGACAGCTGATGACTATTTGCTGGCTATAGCGCTAAATATTCAT 8144
OY 1621 GGTGTATGAGACAGCTATTTCAACCGTGTGACACCACTTTGTAAGCAATGGCTTACGGT 1680
    |||
DB 8143 GGTGTATGAGACAGCTATTTCAACCGTGTGACACCACTTTGTAAGCAATGGCTTACGGT 8084
OY 1681 CGCTGCCAAGTTTACCGCTATGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
    |||
DB 8083 CGCTGCCAAGTTTACCGCTATGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8024
OY 1741 TTGGAATATACAGAGAGGCTCTACAGAAATGAGAGATTCGATTTTAAAAATGGTCTCG 1800
    |||
DB 8023 TTGGAATATACAGAGAGGCTCTACAGAAATGAGAGATTCGATTTTAAAAATGGTCTCG 7964
OY 1801 TTCTACGTGGAAGTCACTGCTCTCAACAAACACCCCATCACTCAAGTCAAGCTCATC 1860
    |||
DB 7963 TTCTACGTGGAAGTCACTGCTCTCAACAAACACCCCATCACTCAAGTCAAGCTCATC 7904
OY 1861 ATCAGATAGTCAACTTCAAGTCTAGCTCAACCACTCCAGAGCAAAATATATGATGACGAC 1920
    |||
DB 7903 ATCAGATAGTCAACTTCAAGTCTAGCTCAACCACTCCAGAGCAAAATATATGATGACGAC 7844
OY 1921 TACCAATCCTAATATATATGAGCAACATCAATATACACCCCTGATCAACAAATTCAGAA 1980
    |||
DB 7843 TACCAATCCTAATATATATGAGCAACATCAATATACACCCCTGATCAACAAATTCAGAA 7784
OY 1981 TCCTCAACGACCAACCA 1999
    |||
DB 7783 TCCTCAACGACCAACCA 7765

```

RESULT 5

US-08-961-527-145/C

Sequence 145, Application US/08961527

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961.527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36, 373

REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 145:

SEQUENCE CHARACTERISTICS:

LENGTH: 10711 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-145

Query Match 100.0%; Score 1999; DB 13; Length 10711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATTAATCACTATCTGCTGACTTGGGTTCTGACCGCGCTCA 60
DB 9767 TAAATCTAGACAAATTAATCACTATCTGCTGACTTGGGTTCTGACCGCGCTCA 9708
QY 61 TGGCAAGCTAATGATATCCACAGATTTGTTAAGGCAATGTTCTATGAGACCA 120
DB 9707 TGGCAAGCTAATGATATCCACAGATTTGTTAAGGCAATGTTCTATGAGACCA 9648
QY 121 TCGCTTCTTCGACACAGGGGGATGATATACCATCCGATCTGGAGCTTCTTGGCCAA 180
DB 9647 TCGCTTCTTCGACACAGGGGGATGATATACCATCCGATCTGGAGCTTCTTGGCCAA 9588
QY 181 TCTGCAAGCAATTCCTCCAGGTGATCACTTCACCCAACTGTTAATGTTGAC 240
DB 9587 TCTGCAAGCAATTCCTCCAGGTGATCACTTCACCCAACTGTTAATGTTGAC 9528
QY 241 TTAATTTCACTTCGACTTCGACCAAGCTATTTCTGTAAGGCTCAGGAAGTTGTT 300
DB 9527 TTAATTTCACTTCGACTTCGACCAAGCTATTTCTGTAAGGCTCAGGAAGTTGTT 9468
QY 301 AGCGATTCAGTTAGAACAAAAGCAACCAAGAAATCTTGACCTATATATAATA 360
DB 9467 AGCGATTCAGTTAGAACAAAAGCAACCAAGAAATCTTGACCTATATATAATA 9408
QY 361 GGTCTACATGTCTAATGGGCACTATGATGACAGCGCTCAAACTATGTTAA 420
DB 9407 GGTCTACATGTCTAATGGGCACTATGATGACAGCGCTCAAACTATGTTAA 9348
QY 421 AGACCTCAATATTTAAGTTTACCTCAGTAGGCTTGGTGGATGGCTCAGGACC 480
DB 9347 AGACCTCAATATTTAAGTTTACCTCAGTAGGCTTGGTGGATGGCTCAGGACC 9288
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGGACCCCAAGCGCGAACTTGTCTT 540
DB 9287 AAACCAATATGACCCCTATTCACATCCAGAGGACCCCAAGCGCGAACTTGTCTT 9228
QY 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTATAC 600
DB 9227 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAACAGTATGAGAAAGCACTATAC 9168
QY 601 ACCAATTAAGTATGAAAGTCAAAATGAGCAAGTAATTAACCTGCTACATGGA 660
DB 9167 ACCAATTAAGTATGAAAGTCAAAATGAGCAAGTAATTAACCTGCTACATGGA 9108
QY 661 TAATTAAGTATGAAAGTCAAAATGAGCAAGTATTAACCTGCTACATGGA 720
DB 9107 TAATTAAGTATGAAAGTCAAAATGAGCAAGTATTAACCTGCTACATGGA 9048
QY 721 AACTGGAGTATGATCTACCAAAATGAGCAAGAGCTCAAAATATGTTGGGATAT 780
DB 9047 AACTGGAGTATGATCTACCAAAATGAGCAAGAGCTCAAAATATGTTGGGATAT 8988
QY 781 TTAATTAAGTATGAAAGTCAAAATGAGCAAGTATTAACCTGCTACATGGA 840
DB 8987 TTAATTAAGTATGAAAGTCAAAATGAGCAAGTATTAACCTGCTACATGGA 8328
QY 841 TGTGATGTTTCAAGGTAAGTATGAGCAAGAGCTCAAAATATGTTGGGATAT 900
DB 8927 TGTGATGTTTCAAGGTAAGTATGAGCAAGAGCTCAAAATATGTTGGGATAT 8868
QY 901 TGTGATGTTTCAAGGTAAGTATGAGCAAGAGCTCAAAATATGTTGGGATAT 960
DB 8867 TGTGATGTTTCAAGGTAAGTATGAGCAAGAGCTCAAAATATGTTGGGATAT 8808
QY 961 ACCGATCAGAGCTATGCTCTGCTGAGTACGAGTATGATGATCACTGCTACTAT 1020
DB 8807 ACCGATCAGAGCTATGCTCTGCTGAGTACGAGTATGATGATCACTGCTACTAT 8748

QY 1021 CGTTCAGATGAGCCCTATTAACCTGCTGGACAAATATCTGTTTATTAACCTGATAG 1080
DB 8747 CGTTCAGATGAGCCCTATTAACCTGCTGGACAAATATCTGTTTATTAACCTGATAG 8688
QY 1081 GGGCTACTTTGGCAACATCACTGTTGCAATAGGCTTCGCAACATCGGAACGCTCAGC 1140
DB 8687 GGGCTACTTTGGCAACATCACTGTTGCAATAGGCTTCGCAACATCGGAACGCTCAGC 8628
QY 1141 CGTGAACCTCTAATCAAGGTGAGCTCAACCGCGCAACACTTCTTAATGTTGTTAG 1200
DB 8627 CGTGAACCTCTAATCAAGGTGAGCTCAACCGCGCAACACTTCTTAATGTTGTTAG 8568
QY 1201 AATCGACTACCAAGTATTCATCTCAATATGCCATTTCAAGTAACCAACCAATGGA 1260
DB 8567 AATCGACTACCAAGTATTCATCTCAATATGCCATTTCAAGTAACCAACCAATGGA 8508
QY 1261 CAAAAATATGAGCAAGTATGTAAGAAAGATGGCTGCTGCTAGCCTGCTTGCMAATG 1320
DB 8507 CAAAAATATGAGCAAGTATGTAAGAAAGATGGCTGCTGCTAGCCTGCTTGCMAATG 8448
QY 1321 TGGAACTTACTATTAACCAATGATATATCAATTAAGTCTGTTAGTGGAGTGAAA 1380
DB 8447 TGGAACTTACTATTAACCAATGATATATCAATTAAGTCTGTTAGTGGAGTGAAA 8388
QY 1381 AGAGTTCTTAATGTGGAAGCTGCTGATGAAGAAACAGACAGCTATATGATGACCA 1440
DB 8387 AGAGTTCTTAATGTGGAAGCTGCTGATGAAGAAACAGACAGCTATATGATGACCA 8328
QY 1441 CATGATGAAAAAGCTCTGACTATGGAAGTGAAGAAATGCTATGCTTGGCTGCC 1500
DB 8327 CATGATGAAAAAGCTCTGACTATGGAAGTGAAGAAATGCTATGCTTGGCTGCC 8268
QY 1501 TCAAGCTGTAAAAAGAGACCTCTAATCTATACAGAGCAAGAAATGGAACCAATCA 1560
DB 8267 TCAAGCTGTAAAAAGAGACCTCTAATCTATACAGAGCAAGAAATGGAACCAATCA 8208
QY 1561 GACCTCTCAATTTGTAGACCTGATGAAGTATTTGCTGCTATACGGTAAATATTCAT 1620
DB 8207 GACCTCTCAATTTGTAGACCTGATGAAGTATTTGCTGCTATACGGTAAATATTCAT 8148
QY 1621 GGTGTATGAGAGGCTATTTCTAAGCTGTGACACCACTTGTAGGCAATGGCTTACGT 1680
DB 8147 GGTGTATGAGAGGCTATTTCTAAGCTGTGACACCACTTGTAGGCAATGGCTTACGT 8088
QY 1681 CGCTGCCAAAGTTTACCGCTCTATGATGACCTGCTGGAAGAAACATCCAGAGA 1740
DB 8087 CGCTGCCAAAGTTTACCGCTCTATGATGACCTGCTGGAAGAAACATCCAGAGA 8028
QY 1741 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGTATTTAAAAATGGTCTG 1800
DB 8027 TTGGAATATACAGAGGGCTCTACAGAAATGAGAAATTCGTATTTAAAAATGGTCTG 7968
QY 1801 TTCTAGTGAAGTCTACCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTATC 1860
DB 7967 TTCTAGTGAAGTCTACCTGCTCCACAAACACCCCATCACTGAAAGTTCAAGCTATC 7908
QY 1861 ATCAGATAGTTCAACTTCAAGTCTACAGCAACAGCTCCAGCACAATATATGATGAC 1920
DB 7907 ATCAGATAGTTCAACTTCAAGTCTACAGCAACAGCTCCAGCACAATATATGATGAC 7848
QY 1921 TACCAATCTTAACATATATGCAACATCAAAATACAAACCTTGATCAACAAATCGAA 1980
DB 7847 TACCAATCTTAACATATATGCAACATCAAAATACAAACCTTGATCAACAAATCGAA 7788
QY 1981 TCTCAACCAAGCAACCA 1999
DB 7787 TCTCAACCAAGCAACCA 7769

RESULT 6
US-09-583-110-1312
; Sequence 1312, Application US/09583110
; GENERAL INFORMATION:

APPLICANT: Lynn Doucelte-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: P47H00-07A
CURRENT APPLICATION NUMBER: US/09/583,110
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 1312
LENGTH: 2160
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
US-09-583-110-1312

Query Match 82.1%; Score 1642; DB 22; Length 2160;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

1 TAAATCTACGACATAAATAATCAATCACTGCTGACTGGGTTCTGAACGCCGCTCA 60
159 taaatctacgacataaataatcaatcaactcattgctgactgggtctgaaagcgcgtcaa 218
61 TGCCCAAGCTAATGATATTCACACAGATTGGTTAAGGCAATGTTTCTATGGAAGACCA 120
219 tgcccaagctaatgatatttcacacagattggttaaggcaatgtttctatcgaaagacca 278
121 TGGCTTTTCGACACAGGGGGATTGATACATCCGTAATCCGGGAGCTTTTCGCGCA 180
279 tggcttttcgacacaggggattgatcatccgtaatccgggagcttttcgcgcaa 338
181 TCTGCAAGCAATTCCTCCACAGGTGATCACTCAACCAAGTTGATTAAAGTTGAC 240
339 tctgcaagcaatttcctccacaggtgatcatcctcaacccaagttgattaaagttgac 398
241 TTACTTTTCACTTGACTTCGACACGACTATTTTCTGTAAGGCTCAGGAAGCTTGGTT 300
399 ttacttttcaacttgacttcgacacgactattttctgtaaggctcaggaagcttggtt 458
301 ACCGATTGCTAGAACAAAAGCAACCAAGACAGAAATCTGACTACTATATAATTA 360
459 accgattgctagaaacaaaagcaacccaagacaagaatcttgactactatataataa 518
361 GGTCTACATGCTAATGGGAATGATGATGACAGACAGCTCAAAATCTACTATGTTAA 420
519 ggtctacatgctaatgggaatgatgatgacagacagctcaaaatctactatagttaa 578
421 AGACTCAATTAATTAAGTTTACCTGAGTTAGCTTGTGCTGGTGAATGCTCAGGACAC 480
579 agactcaatataatgaagtttacctgagtttagcttgcgtggaatgctcaggaacac 638
481 AAACCAATATGACCGCTATTCATCCAGAACGACCAAGCCGACCGGAAATTTGGTCT 540
639 aaaccaaatatgacccgctatttcacccaagaaagcccaagccgcggaacttggctt 698
541 ATCTGAATGAAAAATCAAGGCTACATCTCTGTAACAGTATGAAAAAGCAGTAATAC 600
699 atctgaatgaaaaatcaaggctacatctctgtaaacagtatgaaaaagcagtaataac 758
601 ACCAATTTACTGTGACATCAAAAGTCTCAATCAGCAGTAATTAACCTGCTTACATGGA 660
759 accaatttactgtgacatcaaaagtctcaatcagcagtaatttaacctgcttacaatga 818
661 TAATTAACCTCAAGAGATCAATCAATCAAGTGAAGAAAGAGCTATAAAGCTCTACAC 720
819 taatlaacctcaagagatcaatcaatcaagtgaaagaaagagctataaagctctacac 878
721 AACTGAGATGATGCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGATAT 780

|||||
879 aacttggaatgagtgctacaacaatgtagaccagaagctcaaaaactctgtagatat 938
781 TTACAAATACACAGCAATATACGTTGCCATTCACAGATGAAATTTGCAAGTGCCTTACCAT 840
938 ttacaataacagcagatactgctcctacacagatgatatgcaagtcgctctacacat 998
841 TGTGATGTTTCTAACCGGTAAGTCAATGCCAGTAGAGACGCCATCATCTCAAGTAA 900
998 tgtgatggttctaacggtaaaagtcattgcagctaggaagacgccaatcagtcagtaa 1058
901 TGTTCCTTCGGAATTAACCAAGCAGTAGAAAAACAACCGCATGCGGATCAACTATGAA 960
1059 tgttccttcggaatttaaccaagcagtagaaacaacgcgcagcttgggaaactaatgaa 1118
961 ACCGATACACAGCTATGCTCTGCTTGAGTAGCGTGTCTACGATTCACCTGCTACTAT 1020
1119 accgatacacagctatgctcctgcttggatagcgtatctacatcaactgctactat 1178
1021 CGTTCACGATGAGCCCTATACCTCCGTCGGAACAAATCTCCTGTTTAACTGGGATAG 1080
1179 cgttcacgataagcctataactaccctgggacaataacccctgttataactggatag 1238
1081 GGGCTACTTTGGCAATCACCTTCGCAATACGCCCTGCAACATGCGCAAGCTCCAGC 1140
1239 gggctactttggcaatcaccttcgcaatagcgcctgcaacatactgcgaagctccagc 1298
1141 CGTGAACCTTAACACAGCTGACCTGACACCGCCCAAGACTTCTTAATATGCTTAGG 1200
1299 cgtgaaccttaaacagctgacactgacacgcgcgaagacttctctaaatgcttcgg 1358
1201 AATGACATACCAAGATTAATCAATCAATGCAATTCGATTAACACACCAAGATACGA 1260
1359 aatgacataccaagatataactactcaatgcaatcttaagtaacaacaacgaaatcga 1418
1261 CAAAAAATATGAGACAGTAGTAAGAAAGATGCTGCTGCTTACGCTTTCGCAATAG 1320
1419 caaaaaatatgagacagtagtaagaaagatgctgctgctttagcgttccttgcgaatg 1478
1321 TGGAACTTAATTAACCAATGATATATCATTAAGTGGCTTTAGTGAAGAGTGAAGAA 1380
1479 tggaaacttaataacaaatgataatcaatagctttagtgaagagtgaaagaa 1538
1381 AGAGTTCTATATGCGGAACCTGCGCATGAAGGAAGACAGCCTATATATGACCGA 1440
1539 agagttctatattgcggaacctgcgcatgaaggaagaaagcctataatgatacgga 1598
1441 CATGATGAAGAACAGTCTTGACTTATGAACTGGAAGAAATGCTTATGCTTGCTCC 1500
1599 catgatgaagaacagctctgacttataatgaaactggaagaaatgcttactctgctgc 1658
1501 TGAGCTGTTAAACAGACAGCTTACTATACAGACAGGAAATTTGAAACACCATCA 1560
1659 tgaagctgtttaaacagacagcttactatatacagacggaatgtaaaacacacataa 1718
1561 GACCTCTCAATTTGTAGACCTGATGAATTTGTGCTATACCGCAATTAATTAATCAAT 1620
1719 gacctctcaatttgtagacctgataactatctgctgctataacgctaataatctaac 1778
1621 GGCTGATGACAGGCTATTTCAACGCTGACACACACTTGTAGGCAATGGCTTACGCT 1680
1779 ggcgtgatgacaggctatttcaacgctgacacacacttgaagcaatgagcttaagct 1838
1681 CGCTGCAAAAGTTTACCGCTTATATGATGACCTTACTGCTGTAAGAGAACATCCAGAGA 1740
1839 cgctgcaaaagtttacccgtctatgataccatctatgcttgaaggaagaaatcccaagga 1898
1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGATTTAAATAATGGTCTCG 1800
1899 ttggaatatacagaggggctctacagaaatgagaaattcgattttaaataatggctcgg 1958
1801 TTCTAGTGAAGCTCAGCTGCTCACACACACCCCATCACTGAAAGTTCAAGCTCATC 1860
|||||

Dh 1959 ttctacgtggaactcacctgctccacaacaccccccaatcaactgaagttcaagctcatic 2018
Qy 1861 ATGAGATAGTTCACACTCTAGCTACACCACTCCAGCACAATATAGTAGAC 1920
Dh 2019 atcagatagttcaacttcaagctcagctcaaccaccccaagcaataatagtagac 2078
Qy 1921 TACCAATCTTACAAATATATAGCAACAATCAATACAACTGATCAGACAAATCAGAA 1980
Dh 2079 taccatcccaataataatagcaacaatacaaacctcgtatcaacaataatcagaa 2138
Qy 1981 TCCTCAACGACGACACCA 1999
Dh 2139 tcctcaaccagcaaccca 2157

RESULT 7

US-09-107-433-1102
Sequence 1102, Application US/09107433

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN

THE RAPID TEST

NUMBER OF SEQUENCES: 5206

CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: <Unknown>

OPERATING SYSTEM: <Unknown>

SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,433

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085131

FILING DATE: May 12, 1998

APPLICATION NUMBER: 60/051553

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-011

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1102:

SEQUENCE CHARACTERISTICS:

LENGTH: 2166 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Streptococcus pneumoniae

FEATURE:

NAME/KEY: misc.feature

LOCATION: (B) LOCATION 1..2166

SEQUENCE DESCRIPTION: SEQ ID NO: 1102:

US-09-107-433-1102

Query Match 82.1%; Score 1642; DB 15; Length 2166;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1992; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAAATCTAGACATAAATAATCACTATTGCTGACTGGGTTCTGAGCGCCGCTCAA 60
Dh 165 TAAATCTAGACATAAATAATCACTATTGCTGACTGGGTTCTGAGCGCCGCTCAA 224
Qy 61 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATGCTTTCTATCGAAGCA 120
Dh 225 TGCCCAAGCTAATGATATTTCCACAGATTTGGTTAAGGCAATGCTTTCTATCGAAGCA 284
Qy 121 TCGCTTCTGACACAGGGGGGATTTGATACATCCGATCTGAGGAGCTTTCTGCGCAA 180
Dh 285 TCGCTTCTGACACAGGGGGGATTTGATACATCCGATCTGAGGAGCTTTCTGCGCAA 344
Qy 181 TCTGCAAGCAATTCCTCCAGAGTGATCACTCACCACCAAGTTGATTGATTGATGAC 240
Dh 345 TCTGCAAGCAATTCCTCCAGAGTGATCACTCACCACCAAGTTGATTGATTGATGAC 404
Qy 241 TTTACTTTCAACTTCGACTTCGACGAGCTATTTCTGCTGATAGGCTGAGAGGTTGTT 300
Dh 405 TTTACTTTCAACTTCGACTTCGACGAGCTATTTCTGCTGATAGGCTGAGAGGTTGTT 464
Qy 301 AGGATTCAGTTAGACAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 360
Dh 465 AGGATTCAGTTAGACAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 524
Qy 361 GGTCTACATGCTTAATGGAAGCTATGGAATGACAGACAGCTCAAACTATGTTAA 420
Dh 525 GGTCTACATGCTTAATGGAAGCTATGGAATGACAGACAGCTCAAACTATGTTAA 584
Qy 421 AGACCTCAATATTTAAGTTTACCTGATAGCTTCTGCTGCTGATGCTGCTGATGCTG 480
Dh 585 AGACCTCAATATTTAAGTTTACCTGATAGCTTCTGCTGCTGATGCTGCTGATGCTG 644
Qy 481 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCCAAGAGCTGCTT 540
Dh 645 AAACCAATATGACCCCTATTCACATCCAGAGAGCCCAAGAGCCCAAGAGCTGCTT 704
Qy 541 ATCTGAATGAAAAATCAAGGCTACATCTGCTGTAAGCAAGTATGAGAAAGCAAGTCAAT 600
Dh 705 ATCTGAATGAAAAATCAAGGCTACATCTGCTGTAAGCAAGTATGAGAAAGCAAGTCAAT 764
Qy 601 ACCAATTAAGTATGAGTACAAAGTCAATCAATCAAGTCAAGTCAAGTCAAGTCAAGT 660
Dh 765 ACCAATTAAGTATGAGTACAAAGTCAATCAATCAAGTCAAGTCAAGTCAAGTCAAGT 824
Qy 661 TAAATACCTCAAGAACTCATCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 720
Dh 825 TAAATACCTCAAGAACTCATCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 884
Qy 721 AACTGGATGATGCTTACACAAATGTAGACCAAGAAAGTCAAAACATCTGTGGATAT 780
Dh 885 AACTGGATGATGCTTACACAAATGTAGACCAAGAAAGTCAAAACATCTGTGGATAT 944
Qy 781 TTTCAATACAGAGAAATAGTGGCTATCCAGAGATGGAATTTGCAATCTGCTTACAT 840
Dh 945 TTTCAATACAGAGAAATAGTGGCTATCCAGAGATGGAATTTGCAATCTGCTTACAT 1004
Qy 841 TGTGATGTTTCTAACGGTAAAGTCAATGCGAGCTAGAGAGACGCGCATGCTCAAGTAA 900
Dh 1005 TGTGATGTTTCTAACGGTAAAGTCAATGCGAGCTAGAGAGACGCGCATGCTCAAGTAA 1064
Qy 901 TGTTCCTTGGAAATTAACCAAGCAGTAGAACAACCGGAGCTGGGATCACTATGAA 960
Dh 1065 TGTTCCTTGGAAATTAACCAAGCAGTAGAACAACCGGAGCTGGGATCACTATGAA 1124
Qy 961 ACCGATACAGACTATGCTCTGCTGAGTACAGTATGAGTATGAGTATGAGTATGAGT 1020
Dh 1125 ACCGATACAGACTATGCTCTGCTGAGTACAGTATGAGTATGAGTATGAGTATGAGT 1184
Qy 1021 CGTTACAGATGAGCCCTATATACCTTGAGGACAAATATCTGTTATTAAGTGGATAG 1080
Dh 1185 CGTTACAGATGAGCCCTATATACCTTGAGGACAAATATCTGTTATTAAGTGGATAG 1244
Qy 1081 GGCGTACTTTGGCAATCACTTGAATAGCGCTTGCAATAGCGCTTGCAATAGCGGAAACGTC 1140

```

|||||
Db 1245 GGGCTACTTGGCAACATCATCCTTGCAATACGCCCTGCAACAATGCGGAACGTCCAC 1304
Qy 1141 CGTGGAAATCTTAACAAAGTGGGACTCAACCGCGCAAGACTTTCTTAATGGTCTAG 1200
Db 1305 CGTGGAACTCTTAACAAAGTGGGACTCAACCGCGCAAGACTTTCTTAATGGTCTAG 1364
Qy 1201 AATGCACTACCAAGATTTCACTACTCAATGCCATTTCAAGTAACACAAACGAATCAGA 1260
Db 1365 AATGCACTACCAAGATTTCACTACTCAATGCCATTTCAAGTAACACAAACGAATCAGA 1424
Qy 1261 CAAAAAATATGAGCAAGATAGTGAAGAGATGGCTGCTTACGCTGCCCTTTGCAATGG 1320
Db 1425 CAAAAAATATGAGCAAGATAGTGAAGAGATGGCTGCTTACGCTGCCCTTTGCAATGG 1484
Qy 1321 TGGACTTCTTAATTAACCAATATATATCATATAAGTCGCTTCTTAAGTGAAGTGAANA 1380
Db 1485 TGGAACTTACTATTAACCAATATATATCATATAAGTCGCTTCTTAAGTGAAGTGAANA 1544
Qy 1381 AGAGTTCTTAATGTGCAACTGTCGCACTGATGAAGAAACGACAGCCTATATGATGACGA 1440
Db 1545 AGAGTTCTTAATGTGCAACTGTCGCACTGATGAAGAAACGACAGCCTATATATGATGACGA 1604
Qy 1441 CATGATGAAAACAGTCTTGACTTATGGAATGGAAGAAATGCCCTATCTTGCTGGCTCC 1500
Db 1605 CATGATGAAAACAGTCTTGACTTATGGAATGGAAGAAATGCCCTATCTTGCTGGCTCC 1664
Qy 1501 TTAGGCTGTGTAACAGAACTCTAATATACAGACAGGAAATTTGAACACATCAA 1560
Db 1665 TTAGGCTGTGTAACAGAACTCTAATATACAGACAGGAAATTTGAACACATCAA 1724
Qy 1561 GACCTCTCAATTTGTAGCACTGATGAATATTTGCTGCTTACGCGTAATATTTCAAT 1620
Db 1725 GACCTCTCAATTTGTAGCACTGATGAATATTTGCTGCTTACGCGTAATATTTCAAT 1784
Qy 1621 GGCTGTATGAGAGGCTTATTTACCGCTGACACCACTTGTAGGCAATGGCCTTACGCT 1680
Db 1785 GGCTGTATGAGAGGCTTATTTACCGCTGACACCACTTGTAGGCAATGGCCTTACGCT 1844
Qy 1681 GCGTCGCAAGTTTACCGCTGATGATGACCTGATGATGATGATGATGATGATGATGATG 1740
Db 1845 GCGTCGCAAGTTTACCGCTGATGATGACCTGATGATGATGATGATGATGATGATGATG 1904
Qy 1741 TTGGAATATACAGAGGCTCTTACAGAAATGAGAAATTCGATTTAAATAATGGTCTCG 1800
Db 1905 TTGGAATATACAGAGGCTCTTACAGAAATGAGAAATTCGATTTAAATAATGGTCTCG 1964
Qy 1801 TTCTACGTGGAACCTGCTGCTGCAACACCCCATCACTGAAGTTCAAGCTCATC 1860
Db 1965 TTCTACGTGGAACCTGCTGCTGCAACACCCCATCACTGAAGTTCAAGCTCATC 2024
Qy 1861 ATCAATAGTATTAACCTGCTGCTGCAACACCCCATCACTGAAGTTCAAGCTCATC 1920
Db 2025 ATCAATAGTATTAACCTGCTGCTGCAACACCCCATCACTGAAGTTCAAGCTCATC 2084
Qy 1921 TACCAATCTTAACCAATATATACCAATCAATCAATCAATCAATCAATCAATCAATCA 1980
Db 2085 TACCAATCTTAACCAATATATACCAATCAATCAATCAATCAATCAATCAATCAATCA 2144
Qy 1981 TCCCTACACGACACCA 1999
Db 2145 TCCCTACACGACACCA 2163

RESULT 8
PCT-US02-03987-9325
; SEQUENCE 9325, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
```

```

; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9325
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
; PCT-US02-03987-9325

Query Match 71.9%; Score 1438; DB 1; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1 TAAATCTACGACAAATAAATCACTCATCTGCTGACTGGGTCTGAAAGCGCGCTCAA 60
Db 159 TAAATCTACGACAAATAAATCACTCATCTGCTGACTGGGTCTGAAAGCGCGCTCAA 218
Qy 61 TGCCCAAGCTAATGATATTCACAGATTTGTTAAGCAATCGTTTATGCAAGACCA 120
Db 219 TGCCCAAGCTAATGATATTCACAGATTTGTTAAGCAATCGTTTATGCAAGACCA 278
Qy 121 TCGCTTCTTGACACACAGGAGGATGATACATCCGTAATCTGAGAGCTTTCTGGC 180
Db 279 TCGCTTCTTGACACACAGGAGGATGATACATCCGTAATCTGAGAGCTTTCTGGC 338
Qy 181 TCTGCAAGCAATTCCTCCTCAAGGTGATCAACTCCACCAAGATGATTAATTAAGT 240
Db 339 TCTGCAAGCAATTCCTCCTCAAGGTGATCAACTCCACCAAGATGATTAATTAAGT 398
Qy 241 TTACTTTTCACTTGACCTTCCGACCACTATTTCTGTAAGGCTCAGGAAGCTTG 300
Db 399 TTACTTTTCACTTGACCTTCCGACCACTATTTCTGTAAGGCTCAGGAAGCTTG 458
Qy 301 AGCGATTCAGTATGAAACAAAGCAACCAAGCAAGAAATCTTGACTACTATTAATTA 360
Db 459 AGCGATTCAGTATGAAACAAAGCAACCAAGCAAGAAATCTTGACTACTATTAATTA 518
Qy 361 GGTCTCATGTGTAAATGGGAGATGATGATGATGATGATGATGATGATGATGATGAT 420
Db 519 GGTCTCATGTGTAAATGGGAGATGATGATGATGATGATGATGATGATGATGATGAT 578
Qy 421 AGACCTCAATTAATTAAGTTTACCTGATTAAGCTTGTGGCTGGAATGCCCTGAC 480
Db 579 AGACCTCAATTAATTAAGTTTACCTGATTAAGCTTGTGGCTGGAATGCCCTGAC 638
Qy 481 AAACCAATATGACCCCTATTTACATTCAGACAGACCCCAAGACCCGCAAACTTGT 540
Db 639 AAACCAATATGACCCCTATTTACATTCAGACAGACCCCAAGACCCGCAAACTTGT 698
Qy 541 ATCTGAAATGAATAATCAAGGCTACATCTGCTGTAAGTGAAGAGCAAGTCAATAC 600
Db 699 ATCTGAAATGAATAATCAAGGCTACATCTGCTGTAAGTGAAGAGCAAGTCAATAC 758
Qy 601 ACCAATTAATGATGACCTCAAGTCTCAAAAGTCTCAAAAGCAAGTCAATTAATTA 660
Db 759 ACCAATTAATGATGACCTCAAGTCTCAAAAGTCTCAAAAGCAAGTCAATTAATTA 818
Qy 661 TAATTAATGATGACCTCAAGTCTCAAAAGTCTCAAAAGCAAGTCAATTAATTA 720
Db 819 TAATTAATGATGACCTCAAGTCTCAAAAGTCTCAAAAGCAAGTCAATTAATTA 878
Qy 721 AACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
Db 879 AACTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 938
Qy 781 TTACATACAGACGATACGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 938 TTACATACAGACGATACGATGATGATGATGATGATGATGATGATGATGATGATGAT 996
```


Db 939 ttaacaatagacgaatacgtctgcctatccagacgatgaattgcaagtcgctctaccat 998
QY 841 TGTGATGTTCTTAAGCGTAAAGCATTTGCCCGCTAGAGCGCCGCTAGTCAAGTAA 900
Db 999 tctgaagttctcaacgaagtaacgtaacgtccagctagagcagcagcatgaagtaa 1058
QY 901 TGTTCCTTCGGAATTAACCAAGCAGTAGAAGCAACCGGAGCTGGGATCACTATGAA 960
Db 1059 tgttccctcgtgaattacccaagcagtagaacaacacgcgactgggagatcaatacgaa 1118
QY 961 ACCGATCAGAGACTATGCTCTGCTTGAGTACGCTGTCTAGATCTCAACTGCTACTAT 1020
Db 1119 accgatcagactatgtctcgtcgttgagctacggtgtctacgagcactcgcactat 1178
QY 1021 CGTTACAGAGAGCCCTATTAAGTACCTGGGACAAATAGTCGCTTATTAAGTGGAGTAG 1080
Db 1179 cgttcacgaagagccctataacccctggacaaatacccccgttataacatgaggtag 1238
QY 1081 GGGCTACTTTGGCAACATCACTTGAATACGCCCTGCAAAATCGGAAACGTCGACG 1140
Db 1239 gggctactttggcaacatcaactcgtacatcgccctgcaacatcggaacgtccacg 1298
QY 1141 CGTGGAACTCTTAACAGCTCGGACTCAACCGGCCAAGACTTCTTAATGCTTAGG 1200
Db 1299 cgtgaaaccttaacaaagtcgactcaacgcgcgaagacttccataatgctcgg 1358
QY 1201 AATCGACTACCAAGTATTTACTACTCAATGGCATTTCAAGTAAACCAACGCAATGAG 1260
Db 1359 aatcgactaccaagtaattcaactcaaatgcacttcaagtaaacacacgaatcaga 1418
QY 1261 CAAAAAATATGAGCAAGTAGTAAAAAGATGGCTGCTTACGCTGCTTTGCAAAATG 1320
Db 1419 caaaaaatattgagcgaagtagtgaagaatgctgctcgtctacgctccttgcaaatg 1478
QY 1321 TGGAATCTTATTAATTAACCATGTATATCCATTAAGTGTCTTATAGTGGAGTGA 1380
Db 1479 tggaaacttactataacccaatgtaataccataaagtcgtctttagtgaagtgagaa 1538
QY 1381 AGAGTTCTTAATGTCGGAAGCTGTCGCATGAAGAAACAGACGCTATGTATGACGA 1440
Db 1539 agagttctctaaatgctggaactcgttcgaagaagaacacgcctatatgtatgacgga 1598
QY 1441 CATGATGAAAAAGCTTGAATTTGAGACTGAGCAAGAAATGCTTCTTGGCTGCC 1500
Db 1599 catgataaaaaagctctgagttatgaaactggaacgaatgctatcttctgctccc 1658
QY 1501 TCAGGCTGGTAAACAGGACCTTAACCTATACAGACGAGAAATGCAAAACCAACAA 1560
Db 1659 tcaaggtcgtgtaaaaacaggaaccttaactatacagacgaagaaatgtaaaacacatcaa 1718
QY 1561 GACCTCTCAATTTGTAGACCTGATGAATTAATTTGTGCTATACGCTAAATATTCAAT 1620
Db 1719 gacctctcaattgttagaacctgataactatctgtcgtgataacggtataatattcaat 1778
QY 1621 GGCTGATGAGCAGGCTATTCTTAACCGTCTGACACCACTTGTAGCAATGGCTTACGT 1680
Db 1779 ggtcgtatgagacggctattcttaacgctgacacacactgttagcgaatggccttaccgt 1838
QY 1681 CGGTGCAAGTTTACCGCTCTATGATGACTCTGCTGGAAGGAAATCCAGAGAA 1740
Db 1839 cgttcgcaaaagtttaccgctctatgactacactcgtctgaaaggaagcaatccagagga 1898
QY 1741 TTGGAATTAACAGAGGGGCTTACAGAAATGAGAAATTTGTAATTAATTAATGTTGCTG 1800
Db 1899 ttggaatataccgaagggctctcaacgaataatgagaattcgtatttcaaaaatggtctcg 1958
QY 1801 TTCTACGTGGAAGCTCACTGCTCTCACAAACACCCCATCACTCAAGTGAAGTCAAGCTCATC 1860
Db 1959 ttctacgttgagagctcactcgtctcacacaaccccccatcaactgaagttcaagctcactc 2018
QY 1861 ATCAGATAGTTCAATTTCAACAGTCTAGCTCAACCACTCCAGAGCAAAATATAGTACGAC 1920
Db 2019 atcagatagttcaacttcaacagttcgttcaaccactccaagcacaataatagtagtac 2078

QY 1921 TACCAATCTTAACATATATAGCAACATCAATCAACCCCTGATCAACAAATGAGAA 1980
Db 2079 taccaatcttaacataataatcgcacaatacacaatacacaacccctgatacaaaaatcagaa 2138
QY 1981 TCCCTCAACGACACACCA 1999
Db 2139 tccctcaacgcagcaaacca 2157

RESULT 9
US-09-815-242-9325
; Sequence 9325, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9325
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
; US-09-815-242-9325

Query Match 71.9%; Score 1438; DB 31; Length 2160;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1988; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 TAAATCTAGACAAATTAATTAACATCACTGCTGATGGTTCGTAAGCGCGGCA 60
Db 159 taaatctagacaataataaatacaactcattgctgacttggcttgaacgcgcgtcaa 218
QY 61 TGCCCAAGCTAATGATTTCCACAGATTTGTTAAGCAATGTTCTATTGCAAGACA 120
Db 219 tgcccaagctaatgatatctccacagatttggttaaggaatcgtttctatcgaaagca 278
QY 121 TGCGTTCTTGACACAGGGGATTTGATACATCCGATCTCTGGAGCTTTCTTGCGCAA 180
Db 279 tcgcttcttgacacaggggattgataccatcgtatccttggagcttcttctcgcaa 338
QY 181 TCTGCAAAAGCAATTCCTCAAGGTGATCACTTCACCCACAGTTGATTAAGTTGAC 240
Db 339 tctgcaaaagaattccctcaagtgatcagcttcaactcaacagttgattgaattgac 398

Oy 241 TTACTTTTCACTTGACTTCGACACAGACTATTTCTGTGAAGGCTCAGGAAGCTTGCTT 300
 Db 339 ttactttcaacttcgacttcgaccagactatcttcgtaaggtccaggaagcttgctt 458
 Oy 301 AGCGATTGAGTTAGAACAAAAGCAACAGAACAAATCTTGACTTACTATATAATAA 360
 Db 459 agcgattcagttcagaacaaaagcaacagaaagaaatcttgacctactataataaa 518
 Oy 361 GGTTCACATGCTTATGGAACTATGCAATGAGACAGACAGCTCAAACTACTATGTAA 420
 Db 519 ggttcacatgcttactatggaaactatgaaatgcagacagactcaaaactactatgtaa 578
 Oy 421 AGACCTCAATAATTTAGTTTACCTCAGTTAGCTTGCTGCTGGTAATGCTCAGGCACC 480
 Db 579 agacctcaataatttaagtttaccctcagtttagccttgctggtgtaagcctcaggaacc 638
 Oy 481 AAACCAATATGACCCCTTATTCATCCAGAGACAGCCCAAGCCGCCAATTTGCTT 540
 Db 639 aaaccaataatgaccccttattccatccagaagcagcccaagccgcgaaacttgctt 698
 Oy 541 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGACAGAGTATGAAAAAGCATATAC 600
 Db 699 atctgaatgaaaaatcaaggtactatctctgctgaacagttatggaagcaagctaac 758
 Oy 601 ACCAATTAAGTGTGAGTACAAAAGTCTCAATCAGCAAGTAATTAACCTGCTTACATGA 660
 Db 759 accaatatagtggtgacaaagttccaaatcagcaagtaattaccctgcttaccatga 818
 Oy 661 TAATTAAGTCAAGAGATCATCAATCAAGTTGAAGAACAGAGCTTAACCTACTCAC 720
 Db 819 taattacgtcaaggaagatcatcaatcaagttgaagaagaacaggtctlaacctctac 878
 Oy 721 AACTGGATGATGCTCTACAAATGTAGACCAAGAGAGCTCAAAACATCTGTGGATAT 780
 Db 879 aactggtatgagtgctacacaaatgtagaccagaagactcaaaactctgtyggaat 938
 Oy 781 TTACAATCAGACGAATAGTTGCTTATCCAGACAGTGAATTCGAATGCTTACCAT 840
 Db 939 ttacaatacagacgaataagttgcttaccagaagatgaaatgcaagctcttaccat 998
 Oy 841 TGTATATGTTTCTAACGCTAAAGTATGCTCCAGCTAGAGAGCCCATCAATCAAGTA 900
 Db 999 tgtatgttcttcaacggtlaaagttatgcccagctaggaagacgcctcagltcaagtaa 1058
 Oy 901 TGTTCCTTCGGAATTAACGAAGTAGAAACAAACCGCGATCGGGATCAATATGA 960
 Db 1059 tgttccttcggaattaaccaagcagtagaaacacgcgagctggyggtatcaactgaa 1118
 Oy 961 ACCGATCAGACACTATGCTCTGCTGGAGTACGCTGTCTACGATTCAGTCTACTAT 1020
 Db 1119 accgatcagactatgctctgcttgagtgagtggtgtctacgagtcacatgcacact 1178
 Oy 1021 CGTTACGATGAGCCCTTACTACCTGGGACAAATACTCTGTTTATTAATGGGATAG 1080
 Db 1179 cgttacgagtagcccttactactaccctggaacaataccctgttctlaaactggaatag 1238
 Oy 1081 GGGCTACTTTGGCAACTCAGCTTGCAATACGCCCTGCAACAATGCGGAAGCTCCAGC 1140
 Db 1239 gggctactttggcaactcagcttgcaatagccctgcaacaatcggaagagctccagc 1298
 Oy 1141 CGTGAAGACTTAACAGAGTGAGCTCAACCGCGCAAGACTTCTTAATGTGTAGG 1200
 Db 1299 cgtggaagactttaacagagtgagctcaacccgcgccaagacttcttaaatggtctcg 1358
 Oy 1201 AATGAGTACCAAGTATTCCTACTCAATGCAATTCAGTACAGTACAGTACGATCA 1260
 Db 1359 aatgagtaaccaagtatctactactacaatgcaatctcaagtaacaacaacgaaatcga 1418
 Oy 1261 CAAAAAATATGAGCAATAGTGAAGATGAGCTGCTGCTTACGCTTGTGCAATAG 1320
 Db 1419 caaaaaatatgagcaaatagtgaaagaatgagctgcttcaagctgcttgcgaatag 1478
 Oy 1321 TGGAACTTACTATTAACCAATGTATATCCATTAAGTCTTCTTATGATGGGAGTGA 1380

Db 1479 tggaaactactaaacaaatgtataatcaaaagtgcttctttagtggagtgaaaa 1538
 Oy 1381 AGAGTTCTTATGTCGGAACTCGTCCATGGAAGAAACGACAGCTATATCATGACGA 1440
 Db 1539 agagttcttcaatgctggaaactcggtccatgaaagaagaacagctctataatgaa 1598
 Oy 1441 CATGATGAAAACAGTCTTGACTTATGAACTGAGAAATGCTTCTTGTGCTGCC 1500
 Db 1599 catgataaaaaagctcttgagttatgaaactggaagaaatgcttatcttctgctcc 1658
 Oy 1501 TCAGGCTGTTAAACAGAGACCTTACTATATCAACAGAGAAATTTGAACCATCA 1560
 Db 1659 tcaggctgtlaaaaaagagaccttaactatacagcgggaatgtgaaacacatcaa 1718
 Oy 1561 GACCTCTCAATTTTGAACGCTGATGAACATATTTGCTGCTTACGCGTAATTTCA 1620
 Db 1719 gacctctcaatttgaacacctgaatgaactatcttgcttatacgcgtaaatatcaat 1778
 Oy 1621 GGCTGTATGACAGGCTATTTCTAACCGTCTGACACACTTGTAGCAATGAGCTTAC 1680
 Db 1779 ggcgttatggaagagctatcttaacgcgtctgacacacttgtaggcaatggtcctta 1838
 Oy 1681 CGCTGCCAAAAGTTTACCGCTTATGATGACCTTACTGTCTGAAGAGCAATCCAGAGA 1740
 Db 1839 cgctgccaaaagtttaccgcttataatgacactcgtctgaaggaaagcaatccaga 1898
 Oy 1741 TTGGAATATACAGAGGGGCTCTACGAATGAGAAATTCGATTTAAATAGTGCTCG 1800
 Db 1899 ttggaatatacagaggggctctacagaatgagaaatcgatattgaaatggtgctcg 1958
 Oy 1801 TTCTAGTGGAACTCACCTGCTCCAAACACACCCCATCACTGAAGTTCAAGCTCATC 1860
 Db 1959 ttctagtggaagcttaacctgctccaaacaaccccatcaactgaagtcaagctcac 2018
 Oy 1861 ATCAGATAGTTCACTTCACTGCTTACAGCTTCAACACTCCAGACACAAATATAGTAC 1920
 Db 2019 atcagatagttcaacttcaactcagcttagctcaaccactccaagcacaataatag 2078
 Oy 1921 TACCAATCTTACATTAATACGCAACCAATCAATCAACCTGATTCACAAATATCAGA 1980
 Db 2079 taccaatcttcaataataacgacaacatcaatacaacccttgatcaacaataatgaa 2138
 Oy 1981 TCCTCAACGACACACCA 1999
 Db 2139 tcctcaaccgacacaacca 2157

RESULT 10
 US-10-072-851-9325
 Sequence 9325, Application US/10072851
 GENERAL INFORMATION:
 APPLICANT: Carr, Grant J.
 APPLICANT: Xu, H. Howard
 APPLICANT: Foulkes, J. Gordon
 APPLICANT: Zamudio, Carlos
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlssen, Karl L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Roemer, Terry
 APPLICANT: Jiang, Bo
 APPLICANT: Boone, Charles
 APPLICANT: Bussey, Howard
 TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
 FILE REFERENCE: ELITRA.028A
 CURRENT APPLICATION NUMBER: US/10/072,851
 CURRENT FILING DATE: 2002-02-08
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09

Accession	Sequence	Position
Dd	2079 taaccaatccttaacaataataatcgcaacatcacaataacacccctgtatcaacaanaatcagaa	2138
Qy	1981 TCCTCAACCAAGACACAAACA	1999
Dd	2139 tcctcaaccagcaacaacca	2157

RESULT 11

```

US-60-061-998-596/c
: Sequence 596, Application US/60061998
:
: GENERAL INFORMATION:
:
: APPLICANT: LAGACE, ROBERT E.
: APPLICANT: CORLEY, NEIL C.
: APPLICANT: RUSSO, FRANK D.
: APPLICANT: HANN, AMY L.
: APPLICANT: HEATH, JOE D.
: APPLICANT: FINNEY, GREGORY L.
:
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STEPTOCOCCUS PNEUMONIAE
: TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
:
: NUMBER OF SEQUENCES: 797
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 PORTER DRIVE
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: USA
:
: ZIP: 94304
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/60/061,998
: FILING DATE:
:
: CLASSIFICATION:
:
: ATTORNEY/AGENT INFORMATION:
: NAME: CERRONE, MICHAEL C.
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PM-0006-2P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
:
: INFORMATION FOR SEQ ID NO: 596:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1033 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: MOLECULE TYPE: cDNA
:
: IMMEDIATE SOURCE:
: CLONE: SPN1601
:
: US-60-061-998-596

```

Query Match	71.9%	Score 1438;	DB 45;	Length 10333;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 1988; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;

QY	1	TAAATTCAGCAATATAAAATCAACATCTCTGACTTGGTCTGTGAAGCCCGCTCA	60
Db	9718	TAAATTCAGCAATATAAAATCACTCTTCTGACTTGGTCTGTGAAGCCCGCTCA	9659
QY	61	TGCCAAGCTAATGATATTTCCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACA	120
Db	9658	TGCCAAGCTAATGATATTTCCACAGATTGGTTAAGCAATCGTTTCTATCGAAGACA	9599
QY	121	TGCGTTTTCGACACACAGGGGGATGATATCCATCCGTAATCCGGGAGTTCCTTGCGCA	180
Db	9598	TGCGTTTTCGACACACAGGGGGATGATATCCATCCGTAATCCGGGAGTTCCTTGCGCA	9539
QY	181	TCTGCAAGCAATTTCCCTCCAGAGTGATCAACTCACCACAGTTGATTAAAGTTGAC	240

Dh	9538	TCGCAAAAGCAATTCCTCCAGAGTGGATACGCTCAGCTCACTCAACAGGTTGATTAAGTTGAC	94179
Oy	241	TTACTTTTCAACTTCGACTTCGACCAAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT	300
Dh	9478	TTACTTTTCAAACTTCGACTTCGACCAAGACTATTTCTGTAAGGCTCAGGAAGCTTGGTT	9419
Oy	301	AGGCATTTCAGTTGGAACAAAAAGCAACGAAGCAAGAAATCTTGACCTCACTATATATAA	360
Dh	9418	AGGCATTTCAGTTGGAACAAAAAGCAACGAAGCAAGAAATCTTGACCTCACTATATATAA	9339
Oy	361	GGCTACATGCTTAATGGGAACATATGGAATGCGACAGCAGCTCAAAACTACTATATGTA	420
Dh	9358	GGCTACATGCTTAATGGGAACATATGGAATGCGACAGCAGCTCAAAACTACTATATGTA	9299
Oy	421	AGACCTCAATATTTTAAGTTTACCCTAGTTAGCTTGCCTGGTGAATGCCAGGACACC	480
Dh	9298	AGACCTCAATATTTTAAGTTTACCCTAGTTAGCTTGCCTGGTGAATGCCAGGACACC	9239
Oy	481	AAACCAATATGACCCCTTATTCACATCCAGAGCAGCCCAAGACCCCGCAAACTTGGTCTT	540
Dh	9238	AAACCAATATGACCCCTTATTCACATCCAGAGCAGCCCAAGACCCCGCAAACTTGGTCTT	9179
Oy	541	ATCTGAAATGAAAAATCAGGCTACATCTCTGCTGAACGTTTGAGAAAGCAGTCAATAC	600
Dh	9178	ATCTGAAATGAAAAATCAGGCTACATCTCTGCTGAACGTTTGAGAAAGCAGTCAATAC	9119
Oy	601	ACCAATTAAGTATGAGCTATCAAAAGTCMAAATGACCAATATTTACCCGCTTATCATGA	660
Dh	9118	ACCAATTAAGTATGAGCTATCAAAAGTCTCAAAATGACCAATATTTACCCGCTTATCATGA	9059
Oy	661	TAATTAACCTCAAGGAAGCTCATCAATCAAGTTGAGAAGAAACAGGCTATTAACCTACTAC	720
Dh	9058	TAATTAACCTCAAGGAAGCTCATCAATCAAGTTGAGAAGAAACAGGCTATTAACCTACTAC	8999
Oy	721	AACCTGGATGGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	780
Dh	8998	AACCTGGATGGATGTCTACACAAATGTAGACCAAGAGCTCAAAAACATCTGTGGGATAT	8939
Oy	781	TTACAATACAGACGAATACCTTGCCTATCCAGACGATGTAATGCAAGTCGCTTACCAT	840
Dh	8938	TTACAATACAGACGAATACCTTGCCTATCCAGACGATGTAATGCAAGTCGCTTACCAT	8879
Oy	841	TGTTGATGTTTCTACAGGTAAAGTCAATTGCCACGCTAGSAGCAGCCATCACTCAAGTA	900
Dh	8878	TGTTGATGTTTCTACAGGTAAAGTCAATTGCCACGCTAGSAGCAGCCATCACTCAAGTA	8819
Oy	901	TGTTCCCTGGGAATTTAACCAAGCAGTAGAAGCAAAACGCGACTGGGATCAACTATGAA	960
Dh	8818	TGTTCCCTGGGAATTTAACCAAGCAGTAGAAGCAAAACGCGACTGGGATCAACTATGAA	8759
Oy	961	ACCGATCACAGACTATGCTCTGCTTGGAGTAGAGGTGTACAGATCAACTGCTACTAT	1020
Dh	8758	ACCGATCACAGACTATGCTCTGCTTGGAGTAGAGGTGTACAGATCAACTGCTACTAT	8699
Oy	1021	CGTTACAGATGAGGCCATATACATACCTGGGACAAATACCTCGTTTAACTGGGATAG	1080
Dh	8698	CGTTACAGATGAGGCCATATACATACCTGGGACAAATACCTCGTTTAACTGGGATAG	8639
Oy	1081	GGGCTACTTTGGCAACATCACTTGGCAATATAGCGCTGCACAACTGCGCAAAAGTCCAGC	1140
Dh	8638	GGGCTACTTTGGCAACATCACTTGGCAATATAGCGCTGCACAACTGCGCAAAAGTCCAGC	8579
Oy	1141	CGTGGAAATCTTAACCAAGTCCGAACTCAACCGCGCAAGACTTTTCTTAATGTTGATAGG	1200
Dh	8578	CGTGGAAATCTTAACCAAGTCCGAACTCAACCGCGCAAGACTTTTCTTAATGTTGATAGG	8519
Oy	1201	AATGACATACCACAAGTTTACATACATCAATATGCAATTTAACTAATACACACGAAATCGA	1260
Dh	8518	AATGACATACCACAAGTTTACATACATCAATATGCAATTTAACTAATACACACGAAATCGA	8459
Oy	1261	CAAAAAATATGAGACAGTAGTGAAGAAAGATGGCTGCTTACGCTGCTTTGCAAAATGG	1320
Dh	8458	CAAAAAATATGAGACAGTAGTGAAGAAAGATGGCTGCTTACGCTGCTTTGCAAAATGG	8399


```

Db 552 ATCTGAATGAAAAATCAAGGCTACATCTCTGCTGAAGATGAGAAAGCATCAATAC 611
Qy 601 ACCAATTACTGATGACGATCAAAAGTCTCAATATCAGCAAGTAAATACCTGCTTACATGA 660
Db 612 ACCAATTACTGATGACGATCAAAAGTCTCAATATCAGCAAGTAAATACCTGCTTACATGA 671
Qy 661 TAATTAACGTCGAAGAGTATCATCAATCAAGTTGAAGAAGAGGCTATTAACCTACTCAC 720
Db 672 TAATTAACGTCGAAGAGTATCATCAATCAAGTTGAAGAAGAGGCTATTAACCTACTCAC 731
Qy 721 AACTGGAGATGATGCTTACACAAATGTAGACCAAGAGTCAAAAAATCTGTGGATAT 780
Db 732 AACTGGAGATGATGCTTACACAAATGTAGACCAAGAGTCAAAAAATCTGTGGATAT 791
Qy 781 TTACAATACAGCAATATAGCTTACGCTATCCAGACGATGATGCAAGTGGCTTCTACAT 840
Db 792 TTACAATACAGCAATATAGCTTACGCTATCCAGACGATGATGCAAGTGGCTTCTACAT 851
Qy 841 TGTGATGTTTCTAACGCTAAAGTCAATGCGCAGTAGAGAGAGCCATCACTCAAGTAA 900
Db 852 TGTGATGTTTCTAACGCTAAAGTCAATGCGCAGTAGAGAGAGCCATCACTCAAGTAA 911
Qy 901 TGTTCCTTCCGAATTAACCAAGCAGTAGAACAACCCGCACTGGGGA 949
Db 912 TGTTCCTTCCGAATTAACCAAGCAGTAGAACAACCCGCACTGGGGA 960

```

RESULT 13

PCT-US97-14436-139

Sequence 139, Application PC/TUS9714436

GENERAL INFORMATION:

APPLICANT: Black, Michael

APPLICANT: Hodgson, John

APPLICANT: Knowles, David

APPLICANT: Lonetto, Michael

APPLICANT: Nichols, Richard

TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,

TITLE OF INVENTION: POLYPEPTIDES AND THEIR USES

NUMBER OF SEQUENCES: 713

CORRESPONDENCE ADDRESS:

ADDRESSEE: Smithkline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/14436

CLASSIFICATION:

FILING DATE: 15-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/024,022

FILING DATE: 16-AUG-1996

ATTORNEY/AGENT INFORMATION:

NAME: Gimm, Edward R

REGISTRATION NUMBER: 38,891

REFERENCE/DOCKET NUMBER: P50533

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-4478

TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 139:

SEQUENCE CHARACTERISTICS:

LENGTH: 2172 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA
PCT-US97-14436-139

Query Match 27.1%; Score 541; DB 1; Length 2172;

Best Local Similarity 99.5%; Pred. No. 1,1e-271;

Matches 811; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

Qy 1186 CCTAAATGGTCTAGCAATGAGTACCAAGTATTCCTACTCAATATGCAATTCAGTTAA 1245
Db 13 CCTAAATGGTCTAGCAATGAGTACCAAGTATTCCTACTCAATATGCAATTCAGTTAA 72
Qy 1246 CACAACCGAATCAGACAAAAAATATGAGACCAAGTATGAAAAGATGCTGCTTACGC 1305
Db 73 CACAACCGAATCAGACAAAAAATATGAGACCAAGTATGAAAAGATGCTGCTTACGC 132
Qy 1306 TGCCTTTGCAAAATGGTGAACCTTACTATTAACCAATGTATTCATTAAGTCTTTAG 1365
Db 133 TGCCTTTGCAAAATGGTGAACCTTACTATTAACCAATGTATTCATTAAGTCTTTAG 192
Qy 1366 TGAATGGAGTGAAGAAAGATTCCTAATGTGGAACTGCTGCATGAGAAACGACAC 1425
Db 193 TGAATGGAGTGAAGAAAGATTCCTAATGTGGAACTGCTGCATGAGAAACGACAC 252
Qy 1426 CTATATGATGACCGACATGATGAAAAACGCTTGAATATGAACTGACG- AAATGCC 1484
Db 253 CTATATGATGACCGACATGATGAAAAACGCTTGAATATGAACTGACGAAATGCC 312
Qy 1485 ATCTTGCTTGGCTCCCTCAGGCTGCTTAAACAGAACTCTAATCTATACAGAGAGAA 1544
Db 313 ATCTTGCTTGGCTCCCTCAGGCTGCTTAAACAGAACTCTAATCTATACAGAGAGAA 372
Qy 1545 TTGAAAACCAATCAAGACCTCTCAATTTGTAGACCTGATGAACTAATTTGGCTTATA 1604
Db 373 TTGAAAACCAATCAAGACCTCTCAATTTGTAGACCTGATGAACTAATTTGGCTTATA 432
Qy 1605 CCGCTAAATATTCATGAGCTGTATGAGACAGGCTATTTCAACGCTGACACCACTTTAG 1664
Db 433 CCGCTAAATATTCATGAGCTGTATGAGACAGGCTATTTCAACGCTGACACCACTTTAG 492
Qy 1665 GCAATGCGCTTACGCTGCTGCCAAAGTTTACCGCTTATGATGACCTTCTGTGAAG 1724
Db 493 GCAATGCGCTTACGCTGCTGCCAAAGTTTACCGCTTATGATGACCTTCTGTGAAG 552
Qy 1725 GAAGCAATCCAGAAATGGAATATACAGAGGCGCTTACAGAAATGGAATTTGTAT 1784
Db 553 GAAGCAATCCAGAAATGGAATATACAGAGGCGCTTACAGAAATGGAATTTGTAT 612
Qy 1785 TTAATAATGCTGCTGCTTACGCTGGAACCTGCTCCACAAACCCCATCAACTG 1844
Db 613 TTAATAATGCTGCTGCTTACGCTGGAACCTGCTCCACAAACCCCATCAACTG 672
Qy 1845 AAAGTTCAAGCTCATCATCATGATGATTTCAACTTACAGCTTACAGCTTACAGCTCA 1904
Db 673 AAAGTTCAAGCTCATCATCATGATGATTTCAACTTACAGCTTACAGCTTACAGCTCA 732
Qy 1905 CAATTAATGATGAGACTACATCCCTAACAATATATGCAACAAATATACACCCCTG 1964
Db 733 CAATTAATGATGAGACTACATCCCTAACAATATATGCAACAAATATACACCCCTG 792
Qy 1965 ATCAACCAAAATCAGATCTCTCAACGACGACAAACA 1999
Db 793 ATCAACCAAAATCAGATCTCTCAACGACGACAAACA 827

```

RESULT 14

US-08-911-503-139

Sequence 139, Application US/08911503

GENERAL INFORMATION:

APPLICANT: Black, Michael

APPLICANT: Hodgson, John

APPLICANT: Knowles, David

APPLICANT: Lonetto, Michael

```

? APPLICANT: Nicholas, Richard
? APPLICANT: Stodola, Robert
? TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
? TITLE OF INVENTION: POLYPEPTIDES AND THEIR USES
? NUMBER OF SEQUENCES: 713
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Smithline Beecham Corporation
? STREET: 709 Swedeland Road
? CITY: King of Prussia
? STATE: PA
? COUNTRY: USA
? ZIP: 19406-0939
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/911,503
? FILING DATE: 15-AUG-1997
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 60/024,022
? FILING DATE: 16-AUG-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Gimmil, Edward R
? REGISTRATION NUMBER: 38,891
? REFERENCE/DOCKET NUMBER: P50533
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 610-270-4478
? TELEFAX: 610-270-5090
? TELEX:
? INFORMATION FOR SEQ ID NO: 139:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2172 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: Genomic DNA
? US-08-911-503-139

Query Match      27.1%; Score 541; DB 13; Length 2172;
Best Local Similarity 99.5%; Pred. No. 1.1e-211;
Matches 811; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1186 CTTAAATGCTAGAGATGACATCCCAAGATTCATCTACTCAATGACATTTCAAGTAA 1245
DB 13 CTTAAATGCTAGAGATGACATCCCAAGATTCATCTACTCAATGACATTTCAAGTAA 72
QY 1246 CACAACCGATTCAGACAAAAAATATGAGCAGATAGTGAAGATGGCTGCTTACGC 1305
DB 73 CACAACCGATTCAGACAAAAAATATGAGCAGATAGTGAAGATGGCTGCTTACGC 132
QY 1306 TGCCTTTGCAATGGTGAAGCTTACTATAAACAATGTATATCCATAAAGTCTTTAG 1365
DB 133 TGCCTTTGCAATGGTGAAGCTTACTATAAACAATGTATATCCATAAAGTCTTTAG 192
QY 1366 TGATGGAGTGAAGAAAGATTTCTATATGTCGGAACCTGTCCTTGAAGAAAGCAGAC 1425
DB 193 TGATGGAGTGAAGAAAGATTTCTATATGTCGGAACCTGTCCTTGAAGAAAGCAGAC 252
QY 1426 CTATATGATACGACATGATGTAAGAAAGTCTTACTATAGAACTGAGC-NAATGCCT 1484
DB 253 CTATATGATACGACATGATGTAAGAAAGTCTTACTATAGAACTGAGCAAAATGCCT 312
QY 1485 ATCTTGGCTTGCCTCCAGAGCTGTGTAAGAAACAGAACTCTAATCTATACAGACAGGAAA 1544
DB 313 ATCTTGGCTTGCCTCCAGAGCTGTGTAAGAAACAGAACTCTAATCTATACAGACAGGAAA 372
QY 1545 TTGAAAACCAATCAGACCTCTCAATTTGTAGACACTGATGAATCTTTGCTGGCTATA 1604
DB 373 TTGAAAACCAATCAGACCTCTCAATTTGTAGACACTGATGAATCTTTGCTGGCTATA 432
```

```

QY 1605 CGCCTAAATATTCATGGCTGTATGACAGGCTATTTCAACCGTCTGACACCACTTGTAG 1664
DB 433 CGCCTAAATATTCATGGCTGTATGACAGGCTATTTCAACCGTCTGACACCACTTGTAG 492
QY 1665 GCAATGGCTTACGGTGCCTGCCAAAGTTTACCGCTCTATGATGACCTGCTGTGAAG 1724
DB 493 GCAATGGCTTACGGTGCCTGCCAAAGTTTACCGCTCTATGATGACCTGCTGTGAAG 552
QY 1725 GAACCAATCCAGAGATTTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTAT 1784
DB 553 GAACCAATCCAGAGATTTGGAATATACAGAGGGGCTCTACAGAAATGGAGAAATTCGTAT 612
QY 1785 TTAATAATGGTCTGCTTCTACAGAGTCCAGACTACCTGCTCCACACACCCCATCACTG 1844
DB 613 TTAATAATGGTCTGCTTCTACAGAGTCCAGACTACCTGCTCCACACACCCCATCACTG 672
QY 1845 AAAGTTCAAGCTCATCATCAGATAGTTCAACTTACAGAGTCAAGTCAAGCTCAAGCA 1904
DB 673 AAAGTTCAAGCTCATCATCAGATAGTTCAACTTACAGAGTCAAGTCAAGCTCAAGCA 732
QY 1905 CAATATATGATGACATCAGATCCTTAACATATATAGCAACATCAATCAACCCCTG 1964
DB 733 CAATATATGATGACATCAGATCCTTAACATATATAGCAACATCAATCAACCCCTG 792
QY 1965 ATCAACAAATCAGATCCTCAACACACACCA 1999
DB 793 ATCAACAAATCAGATCCTCAACACACACCA 827

RESULT 15
US-08-911-503A-139
? Sequence 139, Application US/08911503A
? GENERAL INFORMATION:
? APPLICANT: Black, Michael
? APPLICANT: Hodgson, John
? APPLICANT: Knowles, David
? APPLICANT: Lonetto, Michael
? APPLICANT: Nicholas, Richard
? APPLICANT: Stodola, Robert
? TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
? TITLE OF INVENTION: POLYPEPTIDES AND THEIR USES
? NUMBER OF SEQUENCES: 713
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Smithline Beecham Corporation
? STREET: 709 Swedeland Road
? CITY: King of Prussia
? STATE: PA
? COUNTRY: USA
? ZIP: 19406-0939
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/911,503A
? FILING DATE: 15-AUG-1997
? CLASSIFICATION: 536
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 60/024,022
? FILING DATE: 16-AUG-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Gimmil, Edward R
? REGISTRATION NUMBER: 38,891
? REFERENCE/DOCKET NUMBER: P50533
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 610-270-4478
? TELEFAX: 610-270-5090
? TELEX:
? INFORMATION FOR SEQ ID NO: 139:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2172 base pairs
```

```

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
US-08-911-503A-139

```

```

Query Match      27.1%; Score 541; DB 13; Length 2172;
Best Local Similarity 99.5%; Pred. No. 1.1e-271;
Matches 811; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

```

OY 1186 CCTAAATGCTAGAGTAAGTACGACATCCCAAGTATTCCTACTACTCAATGCAATTCCTCAAGTAA 1245
    |||
DB 13 CCTAAATGCTAGAGTAAGTACGACATCCCAAGTATTCCTACTACTCAATGCAATTCCTCAAGTAA 72
OY 1246 CCAACCCGATCAGACAAAAAATATGAGCAAGTACTGAAAAGATGGCTGCTTACGC 1305
    |||
DB 73 CCAACCCGATCAGACAAAAAATATGAGCAAGTACTGAAAAGATGGCTGCTTACGC 132
OY 1306 TGCCTTTGCAATGCTGAACTTACTATTAACCAATGTATTCATCAATGAAGTGTCTTTAG 1365
    |||
DB 133 TGCCTTTGCAATGCTGAACTTACTATTAACCAATGTATTCATCAATGAAGTGTCTTTAG 192
OY 1366 TGATGGAGTGAAGAAAGAGTCTCTAATGTGCGAATCTGCGCATGAAGAAACGACAGC 1425
    |||
DB 193 TGATGGAGTGAAGAAAGAGTCTCTAATGTGCGAATCTGCGCATGAAGAAACGACAGC 252
OY 1426 CTRATATGATGACCGACATATGAAAAACACTCTTGACTTATGCACTGGACG-AAATGCTT 1484
    |||
DB 253 CTRATATGATGACCGACATATGAAAAACACTCTTGACTTATGCACTGGACG-AAATGCTT 312
OY 1485 ATCTTCTGCTGCTCCTCTCAAGGCTGTAAACAGAACTCTACTACTATACAGAGAGAA 1544
    |||
DB 313 ATCTTCTGCTGCTCCTCTCAAGGCTGTAAACAGAACTCTACTACTATACAGAGAGAA 372
OY 1545 TTGAAAAACCATCAAGACACTCTCAATTTGTAGCACCTGATGAATTTGCTGGCTATA 1604
    |||
DB 373 TTGAAAAACCATCAAGACACTCTCAATTTGTAGCACCTGATGAATTTGCTGGCTATA 432
OY 1605 CCGGTAAATATTCATAGGCTGTATGAGACAGGCTATTTCAACGCTGACACCACTTTGAG 1664
    |||
DB 433 CCGGTAAATATTCATAGGCTGTATGAGACAGGCTATTTCAACGCTGACACCACTTTGAG 492
OY 1665 GCAATGCTTACGCTGCTGCTCAAGTATTACGCTCTATGATGACCTTCTGCTGAAG 1724
    |||
DB 493 GCAATGCTTACGCTGCTGCTCAAGTATTACGCTCTATGATGACCTTCTGCTGAAG 552
OY 1725 GAAGCAATCCAGAGATTTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAT 1784
    |||
DB 553 GAAGCAATCCAGAGATTTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAT 612
OY 1785 TTTAAAAATGCTGCTGCTTACGCTGCACTCTCTCAACACACCCCATCACTG 1844
    |||
DB 613 TTTAAAAATGCTGCTGCTTACGCTGCACTCTCTCAACACACCCCATCACTG 672
OY 1845 AAGGTCAAGCTCATATGATAGATTTGAACTTCAACGCTCAACCTCCCAAGAA 1904
    |||
DB 673 AAGGTCAAGCTCATATGATAGATTTGAACTTCAACGCTCAACCTCCCAAGAA 732
OY 1905 CAAATATATGATGACTTACCAATCTTAAACATATATGCAAGCAATCAATTAACACCCCTG 1964
    |||
DB 733 CAAATATATGATGACTTACCAATCTTAAACATATATGCAAGCAATCAATTAACACCCCTG 792
OY 1965 ATCAACAAAATAGAAATCTCTCAACGACGACAAACA 1999
    |||
DB 793 ATCAACAAAATAGAAATCTCTCAACGACGACAAACA 827

```

```

RESULT 16
US-60-068-175-596/c
; Sequence 596, Application US/60068175
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.

```

```

; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; APPLICANT: Finney, Gregory L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF STREPTOCOCCUS PNEUMONIAE
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 1175
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068,175
; FILING DATE: HERewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0009-2 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-416
; INFORMATION FOR SEQ. ID NO: 596:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9845 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: SPN2c598
US-60-068-175-596

```

```

Query Match      14.7%; Score 293; DB 45; Length 9845;
Best Local Similarity 98.9%; Pred. No. 9.3e-142;
Matches 643; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

OY 1 TAAATCTACGACAAATTAATCAACTCACTGCTGACTTGGGTTCTGAACGCGCGCTCAA 60
    |||
DB 9734 TAAATCTACGACAAATTAATCAACTCACTGCTGACTTGGGTTCTGAACGCGCGCTCAA 9675
OY 61 TGGCCAGCTAATGATATTTCCACAGATTTGGTTAAGCAATGTTCTTATGCAAGACCA 120
    |||
DB 9674 TGGCCAGCTAATGATATTTCCACAGATTTGGTTAAGCAATGTTCTTATGCAAGACCA 9615
OY 121 TGGCTCTTCGACACAGGGGGATGATACATCCGATTCGAGAGCTTTCTTGCGCAA 180
    |||
DB 9614 TGGCTCTTCGACACAGGGGGATGATACATCCGATTCGAGAGCTTTCTTGCGCAA 9555
OY 181 TCTGCAACCAATTCCTCCCAAGGTGATCACTCAACCAACAGTATTAAGTTGAC 240
    |||
DB 9554 TCTGCAACCAATTCCTCCCAAGGTGATCACTCAACCAACAGTATTAAGTTGAC 9495
OY 241 TTACTTTTCACTGACTTTCGACAGACTATTTCTGCTAAGGCTCAGAGAGCTTGTT 300
    |||
DB 9494 TTACTTTTCACTGACTTTCGACAGACTATTTCTGCTAAGGCTCAGAGAGCTTGTT 9435
OY 301 AGCGATTGCTTGAAGCAAAAGCAACGACAAAGAAATCTTGACTTACTATATAATTA 360
    |||
DB 9434 AGCGATTGCTTGAAGCAAAAGCAACGACAAAGAAATCTTGACTTACTATATAATTA 9375
OY 361 GGTCTACATGCTAATGGAAGTATGAGATGACAGACAGCTCAAAATCTATATGTTAA 420
    |||

```

Db 9374 GGTCTACATGTCTAATGGAAGTATGGAATGCAGACAGACGCTCAAACTATATGTTAA 9315
QY 421 AGACCTCAATATTTAAGTTAGTTACCTAGTTAGCCTTGCTGGCTGGAATGCTCAGGCACC 480
Db 9314 AGACCTCAATATTTAAGTTAGTTACCTAGTTAGCCTTGCTGGCTGGAATGCTCAGGCACC 9255
QY 481 AAACCAATATGACCCCTATTCACATCCAGAGACGCCAAGACCGCGAAACTTGCTCT 540
Db 9254 AAACCAATATGACCCCTATTCACATCCAGAGACGCCAAGACCGCGAAACTTGCTCT 9195
QY 541 ATCTGAATGAAAAATCAAGGTACATCTCTGCTGGAACATATGAGAAACAGTCATAC 600
Db 9194 ATCTGAATGAAAAATCAAGGTACATCTCTGCTGGAACATATGAGAAACAGTCATAC 9135
QY 601 ACCAATATGATGAGTACAGTCAAGTCAATGAGCAAGTAAATTAACCTG 650
Db 9134 ACCAATATGATGAGTACAGTCAAGTCAATGAGCAAGTAAATTAACCTG 9085

RESULT 17

US-60-029-960-934/C
; Sequence 934, Application US/60029960
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 1649
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/029, 960
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 934:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 323 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-60-029-960-934

Query Match 11.2%; Score 223; DB 41; Length 323;

Best Local Similarity 100.0%; Pred. No. 4,5e-105;

Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1324 AACTTACTATAAACCATATATATCCATAAAGTCGTTTATAGATGGAGGTAAAAAGA 1383
Db 293 AACTTACTATAAACCATATATATCCATAAAGTCGTTTATAGATGGAGGTAAAAAGA 234
QY 1384 GTCTCTAATGTGGAACTGTGCGCATGAAGAAAGACAGCCATATATGATGACCGACAT 1443
Db 233 GTCTCTAATGTGGAACTGTGCGCATGAAGAAAGACAGCCATATATGATGACCGACAT 174

QY 1444 GATGAAAAAGCTCTTACTATATGAACTGAGACGAATAGCTATCTTGGCTCCCTCA 1503
Db 173 GATGAAAAAGCTCTTACTATATGAACTGAGACGAATAGCTATCTTGGCTCCCTCA 114
QY 1504 GGTGTGTAATAACAGAACCTCTACTATACAGACGAGAAATT 1546
Db 113 GGTGTGTAATAACAGAACCTCTACTATATACAGACGAGAAATT 71

RESULT 18

US-09-583-110-1313/C
; Sequence 1313, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PAT000-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 1313
; LENGTH: 204
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-1313

Query Match 10.2%; Score 204; DB 22; Length 204;

Best Local Similarity 100.0%; Pred. No. 4e-95;

Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1762 CTACGAAATGGAATGCTATTTAAAAATGCTGCTGTTCTAGCTGGAATCCTGCTG 1821
Db 204 CTACGAAATGGAATGCTATTTAAAAATGCTGCTGTTCTAGCTGGAATCCTGCTG 145
QY 1822 TCCAGAACACCCCATCACTGAAAGTTCAAGTCTATCATGATGATGTTCAACTTACA 1881
Db 144 TCCAGAACACCCCATCACTGAAAGTTCAAGTCTATCATGATGATGTTCAACTTACA 85
QY 1882 GTCTAGCTCAACCACTCCAGCAATATATAGTACGATCAACCAATCTTAATATATAC 1941
Db 84 GTCTAGCTCAACCACTCCAGCAATATATAGTACGATCAACCAATCTTAATATATAC 25
QY 1942 GCAACAATCAAAATCAACCCCTGA 1965
Db 24 GCAACAATCAAAATCAACCCCTGA 1

RESULT 19

US-09-134-000-1534
; Sequence 1534, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
; FILE REFERENCE: GTC-005
; CURRENT APPLICATION NUMBER: US/09/134,000A
; CURRENT FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 6810
; SEQ ID NO 1534
; LENGTH: 2370
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000-1534

Query Match 1.3%; Score 25; DB 15; Length 2370;

Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1226 TCAATGCCATTTCAGTAACAA 1250
|||||
Db 1495 tcaatgccatttcagtaacaa 1519

RESULT 20
US-09-070-927-217

Sequence 217, Application US/09070927

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

APPLICANT: Patrick J. Dillon

APPLICANT: Steven C. Barash

TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and

NUMBER OF SEQUENCES: 982

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/070,927

FILING DATE: herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: A. Anders Brookes

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB369

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 217:

SEQUENCE CHARACTERISTICS:

LENGTH: 8395 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-09-070-927-217

Query Match

Best Local Similarity 1.3%; Score 25; DB 14; Length 8395;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1226 TCAATGCCATTTCAGTAACAA 1250
|||||
Db 2589 TCAATGCCATTTCAGTAACAA 2613

RESULT 21
US-09-070-927A-217

Sequence 217, Application US/09070927A

GENERAL INFORMATION:

APPLICANT: Charles A. Kunsch

APPLICANT: Patrick J. Dillon

APPLICANT: Steven Barash

TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 982

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/070,927A

FILING DATE: 04-May-2000

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/046,655

FILING DATE: 1997-05-16

APPLICATION NUMBER: 60/044,031

FILING DATE: 1997-05-06

APPLICATION NUMBER: 60/066,009

FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB369

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 217:

SEQUENCE CHARACTERISTICS:

LENGTH: 8395 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 217:

US-09-070-927A-217

Query Match

Best Local Similarity 1.3%; Score 25; DB 14; Length 8395;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1226 TCAATGCCATTTCAGTAACAA 1250
|||||
Db 2589 TCAATGCCATTTCAGTAACAA 2613

RESULT 22
US-09-536-784-227

Sequence 227, Application US/09536784

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997

```
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 227:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-536-784-227
SEQUENCE DESCRIPTION: SEQ ID NO: 227:

Query Match      1.2%; Score 23; DB 20; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 TAAATCTAGACATATAAATC 23
Db      11 TAAATCTAGACATATAAATC 33

RESULT 23
US-09-765-271-227
Sequence 227, Application US/09765271
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,271
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/536,784
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB340P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 227:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-765-271-227
SEQUENCE DESCRIPTION: SEQ ID NO: 227:

Query Match      1.2%; Score 23; DB 30; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-765-271-227
SEQUENCE DESCRIPTION: SEQ ID NO: 227:
```

```
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 TAAATCTAGACATATAAATC 23
Db      11 TAAATCTAGACATATAAATC 33

RESULT 24
US-09-765-272-227
Sequence 227, Application US/09765272
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigenes and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 227:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-765-272-227
SEQUENCE DESCRIPTION: SEQ ID NO: 227:

Query Match      1.2%; Score 23; DB 30; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 TAAATCTAGACATATAAATC 23
Db      11 TAAATCTAGACATATAAATC 33

RESULT 25
US-09-045-649-108
Sequence 108, Application US/60045649
GENERAL INFORMATION:
APPLICANT: Lagace, Robert E.
APPLICANT: Corely, Neil C.
APPLICANT: Russo, Frank D.
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
NUMBER OF SEQUENCES: 1466
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
```

CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/045.649
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PM-0003P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 108:
SEQUENCE CHARACTERISTICS:
LENGTH: 933 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: EF1C0108
US-60-045-649-108

Query Match 1.1%; Score 22; DB 43; Length 933;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1229 AATGCCATTCAAGTAACACAA 1250
|||||
Db 708 AATGCCATTCAAGTAACACAA 729

RESULT 26
US-60-046-653-183
; Sequence 183, Application US/6004653
; GENERAL INFORMATION:
; APPLICANT: Iagace, Robert E.
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/046.653
FILING DATE: HEREMITH
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PM-0003-1 P

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 183:
SEQUENCE CHARACTERISTICS:
LENGTH: 933 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: EF1C183
US-60-046-653-183

Query Match 1.1%; Score 22; DB 43; Length 933;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1229 AATGCCATTCAAGTAACACAA 1250
|||||
Db 708 AATGCCATTCAAGTAACACAA 729

RESULT 27
PCT-US02-03987-6631
; Sequence 6631, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibit
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6631
; LENGTH: 2337
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2337)
PCT-US02-03987-6631

Query Match 1.1%; Score 22; DB 1; Length 2337;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1229 AATGCCATTCAAGTAACACAA 1250
|||||
Db 1465 aatgccattcaagtaacacaa 1486

RESULT 28
US-09-815-242-6631
; Sequence 6631, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6631
; LENGTH: 2337
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2337)
; US-09-815-242-6631
```

```

Query Match          1.1%; Score 22; DB 31; Length 2337;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1229 AATGCCATTTCAGTAACACAA 1250
      |||||||||||||||||||
DB 1465 aatgccatttcagtaacacaa 1486
```

```

RESULT 29
US-10-072-851-6631
; Sequence 6631, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6631
; LENGTH: 2337
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2337)
; US-10-072-851-6631
```

```

Query Match          1.1%; Score 22; DB 37; Length 2337;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1229 AATGCCATTTCAGTAACACAA 1250
      |||||||||||||||||||
DB 1465 aatgccatttcagtaacacaa 1486
```

```

RESULT 30
US-60-068-217-710/C
; Sequence 710, Application US/60068217
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert E.
; APPLICANT: Corley, Neil C.
; APPLICANT: Russo, Frank D.
; APPLICANT: Hann, Amy L.
; APPLICANT: Heath, Joe D.
; APPLICANT: Finney, Gregory L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF ENTEROCOCCUS FAECALIS
; TITLE OF INVENTION: GENOME, FRAGMENTS THEREOF, AND USES THEREOF
; NUMBER OF SEQUENCES: 1239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/068,217
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PM-0003-5 P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 710:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3478 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: genomic DNA
; IMMEDIATE SOURCE:
; CLONE: EFALC710
; US-60-068-217-710
```

```

Query Match          1.1%; Score 22; DB 45; Length 3478;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1229 AATGCCATTTCAGTAACACAA 1250
      |||||||||||||||||||
DB 2771 AATGCCATTTCAGTAACACAA 2750

RESULT 31
US-09-649-163-2312/C
; Sequence 2312, Application US/09649163
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Galvin, Katherine A.
; APPLICANT: Leiby, Kevin R.
```

APPLICANT: Kingsbury, Gillian A.
APPLICANT: Welch, Nadine S.
APPLICANT: McCarthy, Sean A.
APPLICANT: Williamson, Mark
APPLICANT: Richardson, Jennifer
APPLICANT: Macbeth, Kyle J.
APPLICANT: Fraser, Christopher C.
APPLICANT: Villevall, Jean-Luc M.G.
APPLICANT: Goodearl, Andrew D.J.
APPLICANT: Slios-Santiago, Inmaculada
APPLICANT: White, David
APPLICANT: Pan, Yang
APPLICANT: Busfield, Samantha J.
APPLICANT: Deeds, James
APPLICANT: Lee, John
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 1600.1164-001
CURRENT APPLICATION NUMBER: US/09/649,163
CURRENT FILING DATE: 2000-08-25
PRIOR APPLICATION NUMBER: 60/150,608
PRIOR FILING DATE: 1999-08-25
NUMBER OF SEQ ID NOS: 10535
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2312
LENGTH: 537
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(537)
OTHER INFORMATION: n = A,T,C or G
US-09-649-163-2312

Query Match 1.1%; Score 21; DB 25; Length 537;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 572 GCTGAACAGTATGAGAAAGCA 592
|||||
Db 250 GCTGAACAGTATGAGAAAGCA 230

RESULT 32
US-09-652-916-1790/C
Sequence 1790, Application US/09652916
GENERAL INFORMATION:
APPLICANT: Shyjan, Andrew W.
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
TITLE OF INVENTION: THEREFOR
FILE REFERENCE: 1600.1184-001
CURRENT APPLICATION NUMBER: US/09/652,916
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 60/152,103
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 10480
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1790
LENGTH: 537
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(537)
OTHER INFORMATION: n = A,T,C or G
US-09-652-916-1790

Query Match 1.1%; Score 21; DB 25; Length 537;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 572 GCTGAACAGTATGAGAAAGCA 592
|||||
Db 250 GCTGAACAGTATGAGAAAGCA 230

RESULT 33
US-09-403-609A-25
Sequence 25, Application US/09403609A
GENERAL INFORMATION:
APPLICANT: HAKENBECK, Regine
TITLE OF INVENTION: DNA PROBES, METHOD AND KIT FOR IDENTIFYING
TITLE OF INVENTION: ANTIBIOTIC-RESISTANT STRAINS OF BACTERIA
FILE REFERENCE: 012627-011
CURRENT APPLICATION NUMBER: US/09/403,609A
CURRENT FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: PCT/DE98/01134
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: DE 197 17 346.2
PRIOR FILING DATE: 1997-04-24
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 25
LENGTH: 20
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
US-09-403-609A-25

Query Match 1.0%; Score 20; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1577 GCACCTGATGACTATTTC 1596
|||||
Db 1 gcacctgactatttc 20

RESULT 34
US-09-535-896-44152
Sequence 44152, Application US/09535896
GENERAL INFORMATION:
APPLICANT: Selhamer, Jeffrey J.
APPLICANT: Delegeane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullaly, Sara J.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING HYDROLASES
FILE REFERENCE: PD-1003 CIP
CURRENT APPLICATION NUMBER: US/09/535,896
CURRENT FILING DATE: 2000-03-24
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 46268
SOFTWARE: PERL Program
SEQ ID NO 44152
LENGTH: 242
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: hu01334414
US-09-535-896-44152

Query Match 1.0%; Score 20; DB 20; Length 242;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 AAAAGCAACCAAGCAAGAA 338
|||||
Db 36 aaaagcaaccaagcaagaa 55

```
RESULT 35
US-09-515-128-12893/C
; Sequence 12893, Application US/09515128
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandi
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jessen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Ralsi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fel
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/515,128
; CURRENT FILING DATE: 2000-02-28
; EARLIER APPLICATION NUMBER: 09/034,341
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12893
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(332)
; OTHER INFORMATION: n = A,T,C or G
US-09-515-128-12893

Query Match 1.0%; Score 20; DB 19; Length 332;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1867 TACTCACTTCACAGTCTA 1886
      |||||||
Db 243 TACTCACTTCACAGTCTA 224
```

```
RESULT 36
US-08-688-870-855
; Sequence 855, Application US/08688870
; GENERAL INFORMATION:
; APPLICANT: Stuart, Susan G.
; APPLICANT: Delegeane, Angelo
; APPLICANT: Bills, Pamela K.
; APPLICANT: Pham, Mino Thu
; APPLICANT: Altus, Christina M.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Ingrid Erika Akerblom
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
```

```
; TITLE OF INVENTION: PERIPHERAL BLOOD MACROPHAGES
; NUMBER OF SEQUENCES: 2162
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1. for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/688,870
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Michael C. Carione
; REGISTRATION NUMBER: 39132
; REFERENCE/DOCKET NUMBER: PD-0040 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 855:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: 201668
US-08-688-870-855
```

```
Query Match 1.0%; Score 20; DB 10; Length 355;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1769 AATGAGAAATTCGATTATTA 1788
      |||||||
Db 186 AATGAGAAATTCGATTATTA 205
```

```
RESULT 37
US-60-001-754-855
; Sequence 855, Application US/60001754
; GENERAL INFORMATION:
; APPLICANT: Stuart, Susan G.
; APPLICANT: Delegeane, Angelo
; APPLICANT: Bills, Pamela K.
; APPLICANT: Pham, Mino Thu
; APPLICANT: Altus, Christina M.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Ingrid Erika Akerblom
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
; NUMBER OF SEQUENCES: 2162
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 HILVIEW AVENUE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1. for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/60/001,754
```

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: LUTHER, BARBARA J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PD-0040P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0195
TELEFAX: (415) 855-0555
INFORMATION FOR SEQ ID NO: 855:
SEQUENCE CHARACTERISTICS:
LENGTH: 355 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 201668
US-60-001-734-855

Query Match 1.0%; Score 20; DB 39; Length 355;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1769 AATGAGAAATTCGTATTAA 1788
|||||
DB 186 AATGAGAAATTCGTATTAA 205

RESULT 38
US-09-205-070-44712
; Sequence 44712, Application US/09205070
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748
; CURRENT APPLICATION NUMBER: US/09/205,070
; CURRENT FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44712
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-205-070-44712

Query Match 1.0%; Score 20; DB 16; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAAGAA 338
|||||
DB 275 aaaagcaccacgaagaa 294

RESULT 39
US-09-321-214-31801
; Sequence 31801, Application US/09321214
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Drmanac, Radoje
; APPLICANT: Ctkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Dimanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-740

; CURRENT APPLICATION NUMBER: US/09/321,214
; CURRENT FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: US 60/088,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31801
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-321-214-31801

Query Match 1.0%; Score 20; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAAGAA 338
|||||
DB 275 aaaagcaccacgaagaa 294

RESULT 40
US-09-340-623-44712
; Sequence 44712, Application US/09340623
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/340,623
; CURRENT FILING DATE: 1999-06-28
; EARLIER APPLICATION NUMBER: US 09/205,070
; EARLIER FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 44712
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-623-44712

Query Match 1.0%; Score 20; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 AAAAGCAACCAAGCAAGAA 338
|||||
DB 275 aaaagcaccacgaagaa 294

RESULT 41
US-09-516-335-31801
; Sequence 31801, Application US/09516335
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandt
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jensen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroza, Mimi

```

; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisl, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Trian, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/516,335
; EARLIER FILING DATE: 2000-03-01
; EARLIER APPLICATION NUMBER: 09/321,214
; EARLIER FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: 60/088,041
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31801
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-516-335-31801
```

```

Query Match          1.0%; Score 20; DB 19; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 319 AAAAGCAACCAAGCAAGAA 338
Db 275 aaaagcaaccaagcaagaa 294
```

```

RESULT 42
; Sequence 31801, Application US/09733811
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroja, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
```

```

; APPLICANT: Raisl, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/733,811
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/321,214
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31801
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-733-811-31801
```

```

Query Match          1.0%; Score 20; DB 29; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 319 AAAAGCAACCAAGCAAGAA 338
Db 275 aaaagcaaccaagcaagaa 294
```

```

RESULT 43
; Sequence 31801, Application US/09733811A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-740
; CURRENT APPLICATION NUMBER: US/09/733,811A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 09/321,214
; PRIOR FILING DATE: 1999-05-26
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31801
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-733-811A-31801
```

```

Query Match          1.0%; Score 20; DB 29; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 319 AAAAGCAACCAAGCAAGAA 338
Db 275 aaaagcaaccaagcaagaa 294
```

```

RESULT 44
US-09-898-888-44712
```



```

: Sequence 44712, Application US/09698888
:
: GENERAL INFORMATION:
:
: APPLICANT: Hyseq, Inc.
:
: TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
:
: TITLE OR INVENTION: LIBRARIES
:
: FILE REFERENCE: 20411-748CON1
:
: CURRENT APPLICATION NUMBER: US/09/898,888
:
: CURRENT FILING DATE: 2001-07-03
:
: PRIOR APPLICATION NUMBER: 09/340,623
:
: PRIOR FILING DATE: 1999-06-28
:
: NUMBER OF SEQ ID NOS: 45207
:
: SOFTWARE: FastSeq for Windows Version 3.0
:
: SEQ ID NO 44712
:
: LENGTH: 371
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: OS-09-898-888-44712

```

Query Match	1.0%;	Score 20;	DB 33;	Length 371;
Best Local Similarity	100.0%;	Pred. No. 98;		
Matches	20;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0

Qy	319	AAAAGCAACCAAGCAAGAA	338
Db	275	aaaagcaaccaagcaagaa	294

```

RESULT 45
US-09-898-888A-44712
: Sequence 44712. Application US/09898888A
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: TITLE OF INVENTION: NOVEL COMTIGS OBTAINED FROM VARIOUS CDNAN
: TITLE OF INVENTION: LIBRARIES
: FILE REFERENCE: 20411-748CON1
: CURRENT APPLICATION NUMBER: US/09/898,888A
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: US/09/340,623
: PRIOR FILING DATE: 1999-06-28
: PRIOR APPLICATION NUMBER: US 09/205,070
: PRIOR FILING DATE: 1998-12-03
: NUMBER OF SEQ ID NOS: 45207
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 44712
: LENGTH: 371
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-898-888A-44712

```

Query Match	1.0%;	Score 20;	DB 33;	Length 371;
Best Local Similarity	100.0%;	Pred. No. 98;		
Matches	20;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

```

Qy      319 AAAAGCAACCAAGCAAGAA 338
          |||||
Db      275 aaagcaaccaagcaagaa 294

```

RESULT 46
US-09-975-640-31801
Sequence 31801, Application US/09975640
GENERAL INFORMATION:
APPLICANT: Arterburn, Matthew
APPLICANT: Asgnari, Vida
APPLICANT: Damavandi, Simin
APPLICANT: Dickson, Mark
APPLICANT: Drake, Jim
APPLICANT: Drmenac, Radoje
APPLICANT: Engleman, Carrie
APPLICANT: Faulkner, Brandy
APPLICANT: Fox, Melvin

```

APPLICANT: Garcia, Veronica
APPLICANT: Giedt, Gretchen
APPLICANT: Uessen, Aaron
APPLICANT: Jomek, Leni
APPLICANT: Jones, Lee
APPLICANT: Kita, David
APPLICANT: Labat, Ivan
APPLICANT: Laroya, Mimi
APPLICANT: Lomelli, Michelle
APPLICANT: Nelken, Sarah
APPLICANT: Nguyen, Kodý
APPLICANT: Nguyen, Lynne
APPLICANT: Nguyen, Phuong
APPLICANT: Nogara, Margie
APPLICANT: Palencia, Servando
APPLICANT: Ralsi, Fariba
APPLICANT: Randhwa, Gurpreet
APPLICANT: Sidhu, Navjivan
APPLICANT: Smith, Benjamin
APPLICANT: Smythe, Ashleigh
APPLICANT: Tkach, Joe
APPLICANT: Tran, Lien
APPLICANT: Verna, Ron
APPLICANT: Wechter, Adam
APPLICANT: Wu, Kenneth
APPLICANT: Ylm, Kenneth
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 740CJP
CURRENT APPLICATION NUMBER: US/09/975,640
CURRENT FILING DATE: 2001-10-11
PRIOR APPLICATION NUMBER: 09/733,811
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/088,041
PRIOR FILING DATE: 1998-06-02
NUMBER OF SEQ ID NOS: 31906
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31801
LENGTH: 371
TYPE: DNA
ORGANISM: Homo sapiens
IS-09-975-640-31801

```

Query Match	1.0%;	Score 20;	DB 36;	Length 371;
Best Local Similarity	100.0%;	Pred. No. 98;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	319	AAAGCAACCAAGCAAGAA	338
Db	275	aaagcaaccaagcaagaa	294

```

RESULT 47
US-09-975-640A-1801
Sequence 31801, Application US/09975640A
GENERAL INFORMATION:
APPLICANT: Artterburn, Matthew
APPLICANT: Asghari, Vida
APPLICANT: Danaevandi, Slatin
APPLICANT: Dickson, Mark
APPLICANT: Drake, Jim
APPLICANT: Dirmanac, Radoje
APPLICANT: Engleman, Carrie
APPLICANT: Faulkner, Brandy
APPLICANT: Fox, Melvin
APPLICANT: Garcia, Veronica
APPLICANT: Giedt, Gretchen
APPLICANT: Jessen, Aaron
APPLICANT: Jones, Leil
APPLICANT: Jones, Lee
APPLICANT: Kita, David
APPLICANT: Lapat, Ivan

```

```

; APPLICANT: Iaroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kodj
; APPLICANT: Nguyen, Lyne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Ralsi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjivan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640A
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31801
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-975-640A-31801

```

```

Query Match 1.0%; Score 20; DB 36; Length 371;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 319 AAAAGCACCACAGCAGAAA 338
Db 275 aaaagcaaccaagcaagaaa 294

```

```

RESULT 48
US-09-271-490-13695/c
; Sequence 13695, Application US/09271490
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-767
; CURRENT APPLICATION NUMBER: US/09/271,490
; CURRENT FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 19424
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13695
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(378)
; OTHER INFORMATION: n = A,T,C or G
US-09-271-490-13695

```

```

Query Match 1.0%; Score 20; DB 16; Length 378;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1867 TAGTTCAACTTCACAGCTCTA 1886
Db 131 TAGTTCAACTTCACAGCTCTA 112

```

```

RESULT 49
US-09-925-552-13695/c
; Sequence 13695, Application US/0925552
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-767
; CURRENT APPLICATION NUMBER: US/09/925,552
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/271,490
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 19424
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13695
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(378)
; OTHER INFORMATION: n = A,T,C or G
US-09-925-552-13695

```

```

Query Match 1.0%; Score 20; DB 34; Length 378;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1867 TAGTTCAACTTCACAGCTCTA 1886
Db 131 TAGTTCAACTTCACAGCTCTA 112

```

```

RESULT 50
US-10-032-354-13695/c
; Sequence 13695, Application US/10032354
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-767
; CURRENT APPLICATION NUMBER: US/10/032,354
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/271,490
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 19424
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13695
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(378)
; OTHER INFORMATION: n = A,T,C or G
US-10-032-354-13695

```

```

Query Match 1.0%; Score 20; DB 37; Length 378;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1867 TAGTTCAACTTCACAGCTCTA 1886
Db 131 TAGTTCAACTTCACAGCTCTA 112

```

Search completed: June 13, 2002, 13:40:40
Job time: 9593 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 13, 2002, 11:34:23 ; Search time 277.73 Seconds
(without alignments)
10116.359 Million cell updates/sec

Title: US-08-961-083-1
Perfect score: 1999
Sequence: 1 TAAATCTACGACAAATAAAA.....ATCCTCAACGACGACAAACA 1999

Scoring table: OLIGO_NMC
Gapop 60.0 , Gapext 60.0

Searched: 997330 seqs, 70275440 residues

Word size : 15

Total number of hits satisfying chosen parameters: 2617

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCR_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1999	100.0	1999	4	US-08-961-083-1
2	23	1.2	33	4	Sequence 227, App
3	21	1.1	1389	5	US-09-540-209B-4827
4	20	1.0	332	5	US-09-721-544-12893
5	20	1.0	502	6	US-10-027-632-240647
6	20	1.0	502	6	US-10-027-632-240647
7	20	1.0	585	6	US-10-027-632-240648
8	19	1.0	4508	5	US-09-053-375B-301
9	19	1.0	427	5	US-09-620-393B-634
10	19	1.0	408	5	US-09-620-393B-8776
11	19	1.0	507	6	US-10-027-632-250559
12	19	1.0	1079	6	US-10-027-632-250559
13	19	1.0	1079	6	US-10-027-632-250570
14	19	1.0	1079	6	US-10-027-632-250571
15	19	1.0	1299	7	US-60-360-039-23856
16	19	1.0	1299	7	US-60-360-039-24027
17	19	1.0	1920	7	US-60-360-039-42540
18	18	0.9	166	5	US-09-539-331D-25958
19	18	0.9	229	5	US-09-539-800C-2484
20	18	0.9	250	5	US-09-540-210B-30401
21	18	0.9	255	5	US-09-539-331D-33884
22	18	0.9	255	5	US-09-539-800C-15921
23	18	0.9	264	5	US-09-539-800C-16898
24	18	0.9	268	5	US-09-540-210B-24058
25	18	0.9	277	5	US-09-919-002-7641
26	18	0.9	285	5	US-09-539-331D-21402

27	18	0.9	296	5	US-09-539-331D-34826	Sequence 34826, A
28	18	0.9	457	6	US-10-027-632-94603	Sequence 94603, A
29	18	0.9	457	6	US-10-027-632-94604	Sequence 94604, A
30	18	0.9	457	6	US-10-027-632-305498	Sequence 305498, A
31	18	0.9	457	6	US-10-027-632-305499	Sequence 305499, A
32	18	0.9	457	6	US-10-027-632-305500	Sequence 305500, A
33	18	0.9	478	6	US-10-027-632-286371	Sequence 286371, A
34	18	0.9	512	6	US-10-097-105-1069	Sequence 1069, Ap
35	18	0.9	526	5	US-09-620-393B-6548	Sequence 6548, Ap
36	18	0.9	532	6	US-10-027-632-135305	Sequence 135305, A
37	18	0.9	581	6	US-10-027-632-81926	Sequence 81926, A
38	18	0.9	850	6	US-10-027-632-157285	Sequence 157285, A
39	18	0.9	1229	6	US-10-106-698-791	Sequence 791, App
40	18	0.9	1281	5	US-09-935-625-18639	Sequence 18639, A
41	18	0.9	1281	5	US-09-935-625-27058	Sequence 27058, A
42	18	0.9	1293	7	US-60-360-039-47071	Sequence 47071, A
43	18	0.9	1299	7	US-60-360-039-47214	Sequence 47214, A
44	18	0.9	1328	5	US-09-919-002-13066	Sequence 13066, A
45	18	0.9	1555	6	US-10-027-632-253001	Sequence 253001, A
46	18	0.9	1555	6	US-10-027-632-253002	Sequence 253002, A
47	18	0.9	1555	6	US-10-027-632-253003	Sequence 253003, A
48	18	0.9	1555	6	US-10-027-632-253004	Sequence 253004, A
49	18	0.9	1773	7	US-60-360-039-24796	Sequence 24796, A
50	18	0.9	2288	6	US-10-115-123-53	Sequence 53, App1
51	18	0.9	2773	6	US-10-027-632-111835	Sequence 111835, A
52	18	0.9	3201	1	PCR-US02-13142-2190	Sequence 2190, Ap
53	18	0.9	3201	1	PCR-US02-13142-7190	Sequence 7190, Ap
54	18	0.9	3201	6	US-10-128-714-2190	Sequence 2190, Ap
55	18	0.9	3201	6	US-10-128-714-7190	Sequence 7190, Ap
56	18	0.9	3293	1	PCR-US02-13142-1190	Sequence 1190, Ap
57	18	0.9	3293	1	PCR-US02-13142-6190	Sequence 6190, Ap
58	18	0.9	3293	6	US-10-128-714-1190	Sequence 1190, Ap
59	18	0.9	3293	6	US-10-128-714-6190	Sequence 6190, Ap
60	18	0.9	5293	1	PCR-US02-13142-1190	Sequence 1190, App
61	18	0.9	5293	1	PCR-US02-13142-5190	Sequence 5190, Ap
62	18	0.9	5293	6	US-10-128-714-1190	Sequence 1190, App
63	18	0.9	5293	6	US-10-128-714-5190	Sequence 5190, Ap
64	18	0.9	13165	6	US-10-105-299-9019	Sequence 9019, Ap
65	18	0.9	13165	6	US-10-116-355-459	Sequence 459, App
66	18	0.9	22008	1	PCR-US02-09239-242	Sequence 242, App
67	18	0.9	22008	1	PCR-US02-09188-900	Sequence 900, App
68	18	0.9	22008	1	PCR-US02-09257-581	Sequence 581, App
69	18	0.9	22008	1	PCR-US02-09370-930	Sequence 930, App
70	18	0.9	22008	1	PCR-US02-09922-568	Sequence 568, App
71	18	0.9	22008	6	US-10-105-299-6302	Sequence 6302, Ap
72	18	0.9	713059	6	US-10-027-632-174581	Sequence 174581, A
73	17	0.9	27	4	US-08-961-083-228	Sequence 228, App
74	17	0.9	220	5	US-10-027-632-70744	Sequence 70744, A
75	17	0.9	250	5	US-09-539-331D-7293	Sequence 7293, Ap
76	17	0.9	252	5	US-09-975-254-17107	Sequence 17107, A
77	17	0.9	253	5	US-09-975-254-2423	Sequence 2423, Ap
78	17	0.9	253	5	US-09-460-592B-136	Sequence 136, App
79	17	0.9	264	6	US-10-027-632-72513	Sequence 72513, A
80	17	0.9	276	5	US-09-540-210B-34006	Sequence 34006, A
81	17	0.9	382	5	US-09-721-544-13637	Sequence 13637, A
82	17	0.9	385	6	US-10-027-632-39314	Sequence 39314, A
83	17	0.9	402	6	US-10-099-926-3300	Sequence 330, App
84	17	0.9	437	5	US-09-539-331D-29344	Sequence 29344, A
85	17	0.9	443	6	US-10-027-632-189682	Sequence 189682, A
86	17	0.9	457	6	US-10-027-632-79353	Sequence 79353, A
87	17	0.9	457	6	US-10-027-632-79354	Sequence 79354, A
88	17	0.9	457	6	US-10-027-632-301173	Sequence 301173, A
89	17	0.9	457	6	US-10-027-632-301174	Sequence 301174, A
90	17	0.9	472	6	US-10-027-632-88245	Sequence 88245, A
91	17	0.9	546	6	US-10-027-632-238977	Sequence 238977, A
92	17	0.9	546	6	US-10-027-632-238978	Sequence 238978, A
93	17	0.9	558	6	US-10-027-632-89473	Sequence 89473, A
94	17	0.9	558	6	US-10-027-632-89472	Sequence 89472, A
95	17	0.9	558	6	US-10-027-632-89474	Sequence 89474, A
96	17	0.9	565	6	US-10-027-632-262082	Sequence 262082, A
97	17	0.9	573	6	US-10-027-632-26318	Sequence 26318, A
98	17	0.9	623	6	US-10-027-632-3300	Sequence 330, App
99	17	0.9	623	6	US-10-027-632-3345	Sequence 345, App

c 100	17	0.9	633	6	US-10-027-632-54574	Sequence 54574, A	c 173	16	0.8	386	6	US-10-125-968-905	Sequence 905, App
c 101	17	0.9	633	6	US-10-027-632-293998	Sequence 293998, A	c 174	16	0.8	386	6	US-10-027-632-178887	Sequence 178887, A
c 102	17	0.9	642	6	US-10-027-632-256087	Sequence 256087, A	c 175	16	0.8	395	6	US-10-116-712-634	Sequence 634, App
c 103	17	0.9	643	6	US-10-027-632-211646	Sequence 211646, A	c 176	16	0.8	397	6	US-10-099-926-1717	Sequence 1717, App
c 104	17	0.9	643	6	US-10-027-632-211647	Sequence 211647, A	c 177	16	0.8	404	5	US-09-721-544-1213	Sequence 1213, App
c 105	17	0.9	643	6	US-10-027-632-211648	Sequence 211648, A	c 178	16	0.8	425	5	US-09-721-544-22282	Sequence 22282, App
c 106	17	0.9	644	6	US-10-027-632-192473	Sequence 192473, A	c 179	16	0.8	430	6	US-10-027-632-294776	Sequence 294776, A
c 107	17	0.9	662	6	US-10-027-632-127685	Sequence 127685, A	c 180	16	0.8	431	5	US-09-919-002-7411	Sequence 7411, App
c 108	17	0.9	711	6	US-10-027-632-22033	Sequence 22033, A	c 181	16	0.8	432	6	US-10-027-632-308516	Sequence 308516, A
c 109	17	0.9	720	6	US-10-027-632-23633	Sequence 23633, A	c 182	16	0.8	437	6	US-10-027-632-94829	Sequence 94829, A
c 110	17	0.9	730	6	US-10-027-632-23634	Sequence 23634, A	c 183	16	0.8	440	6	US-10-027-632-65703	Sequence 65703, A
c 111	17	0.9	734	6	US-10-027-632-11064	Sequence 11064, A	c 184	16	0.8	460	5	US-09-685-404A-203	Sequence 203, App
c 112	17	0.9	734	6	US-10-027-632-11065	Sequence 11065, A	c 185	16	0.8	462	5	US-09-685-404A-207	Sequence 207, App
c 113	17	0.9	776	6	US-10-027-632-168427	Sequence 168427, A	c 186	16	0.8	464	5	US-09-620-3938-6694	Sequence 6694, App
c 114	17	0.9	779	6	US-10-027-632-145394	Sequence 145394, A	c 187	16	0.8	467	5	US-09-620-3938-6694	Sequence 6694, App
c 115	17	0.9	975	7	US-09-360-039-40220	Sequence 40220, A	c 188	16	0.8	468	6	US-10-027-632-189086	Sequence 189086, A
c 116	17	0.9	1058	5	US-09-919-002-5787	Sequence 5787, App	c 189	16	0.8	468	6	US-10-027-632-189087	Sequence 189087, A
c 117	17	0.9	1202	5	US-09-882-227-509	Sequence 509, App	c 190	16	0.8	470	6	US-10-027-632-84940	Sequence 84940, A
c 118	17	0.9	1211	5	US-09-919-002-1145	Sequence 1145, App	c 191	16	0.8	470	6	US-10-027-632-178926	Sequence 178926, A
c 119	17	0.9	1244	5	US-09-919-002-2529	Sequence 2529, App	c 192	16	0.8	471	6	US-10-027-632-82965	Sequence 82965, A
c 120	17	0.9	1284	7	US-09-360-039-47007	Sequence 47007, A	c 193	16	0.8	471	6	US-10-027-632-82966	Sequence 82966, A
c 121	17	0.9	1457	6	US-10-130-080-3	Sequence 3, App1	c 194	16	0.8	473	5	US-09-620-3938-5469	Sequence 5469, App
c 122	17	0.9	1531	6	US-10-115-123-32	Sequence 32, App1	c 195	16	0.8	476	6	US-10-027-632-50656	Sequence 50656, App
c 123	17	0.9	1906	5	US-09-442-3848-599	Sequence 599, App	c 196	16	0.8	480	6	US-10-027-632-41506	Sequence 41506, App
c 124	17	0.9	2013	6	US-10-106-698-1619	Sequence 1619, App	c 197	16	0.8	484	5	US-09-721-544-5673	Sequence 5673, App
c 125	17	0.9	2163	7	US-09-360-039-36422	Sequence 36422, A	c 198	16	0.8	486	7	US-60-377-240-7841	Sequence 7841, App
c 126	17	0.9	2346	6	US-10-027-632-110029	Sequence 110029, A	c 199	16	0.8	488	5	US-09-721-544-5672	Sequence 5672, App
c 127	17	0.9	2497	7	US-09-360-039-26849	Sequence 26849, A	c 200	16	0.8	492	5	US-09-721-544-2085	Sequence 2085, App
c 128	17	0.9	3247	6	US-10-146-473-34	Sequence 34, App1	c 201	16	0.8	494	6	US-10-027-632-69908	Sequence 69908, A
c 129	17	0.9	3510	7	US-09-360-039-24782	Sequence 24782, A	c 202	16	0.8	507	6	US-10-027-632-34500	Sequence 34500, A
c 130	17	0.9	22680	6	US-10-105-299-11585	Sequence 11585, A	c 203	16	0.8	513	6	US-10-027-632-188432	Sequence 188432, A
c 131	17	0.9	54863	6	US-10-105-299-8039	Sequence 8039, App	c 204	16	0.8	519	6	US-10-027-632-6820	Sequence 6820, App
c 132	17	0.9	54877	1	PCT-US02-09188-1314	Sequence 1314, App	c 205	16	0.8	524	6	US-10-027-632-312273	Sequence 312273, A
c 133	17	0.9	54877	1	PCT-US02-09370-1372	Sequence 1372, App	c 206	16	0.8	527	6	US-10-027-632-6809	Sequence 6809, App
c 134	17	0.9	54877	1	PCT-US02-09922-863	Sequence 863, App	c 207	16	0.8	527	6	US-10-027-632-323318	Sequence 323318, A
c 135	17	0.9	54877	6	US-10-105-299-6896	Sequence 6896, App	c 208	16	0.8	529	5	US-09-721-544-1228	Sequence 1228, App
c 136	17	0.9	54877	6	US-10-105-299-8040	Sequence 8040, App	c 209	16	0.8	530	5	US-09-620-3938-8742	Sequence 8742, App
c 137	17	0.9	12034	6	US-10-105-299-9867	Sequence 9867, App	c 210	16	0.8	532	6	US-10-027-632-58115	Sequence 58115, A
c 138	17	0.9	160820	6	PCT-US02-11086-5	Sequence 5, App1	c 211	16	0.8	542	6	US-10-027-632-59136	Sequence 59136, A
c 139	17	0.9	2940917	6	US-10-027-632-174763	Sequence 174763, A	c 212	16	0.8	549	5	US-09-991-936-1290	Sequence 1290, App
c 140	16	0.8	20	6	US-10-130-080-55	Sequence 55, App1	c 213	16	0.8	550	7	US-60-377-240-5728	Sequence 5728, App
c 141	16	0.8	168	5	US-09-539-331D-17706	Sequence 17706, A	c 214	16	0.8	554	6	US-10-027-632-281424	Sequence 281424, A
c 142	16	0.8	180	5	US-09-539-331D-4063	Sequence 4063, App	c 215	16	0.8	557	6	US-10-027-632-77645	Sequence 77645, A
c 143	16	0.8	188	6	US-10-027-632-53214	Sequence 53214, A	c 216	16	0.8	557	6	US-10-027-632-300549	Sequence 300549, A
c 144	16	0.8	208	5	US-09-539-800C-5010	Sequence 5010, App	c 217	16	0.8	561	6	US-10-027-632-134117	Sequence 134117, A
c 145	16	0.8	216	5	US-09-899-575-80	Sequence 80, App1	c 218	16	0.8	571	6	US-10-027-632-215927	Sequence 215927, A
c 146	16	0.8	239	5	US-09-540-210B-29605	Sequence 29605, A	c 219	16	0.8	572	6	US-10-027-632-71487	Sequence 71487, A
c 147	16	0.8	245	5	US-09-539-331D-21787	Sequence 21787, A	c 220	16	0.8	572	6	US-10-027-632-293800	Sequence 293800, A
c 148	16	0.8	246	5	US-09-539-331D-27254	Sequence 27254, A	c 221	16	0.8	573	6	US-10-027-632-235623	Sequence 235623, A
c 149	16	0.8	247	5	US-09-539-806B-1690	Sequence 1690, App	c 222	16	0.8	577	6	US-10-105-299-9026	Sequence 9026, App
c 150	16	0.8	248	5	US-09-539-800C-9055	Sequence 9055, App	c 223	16	0.8	586	6	US-10-027-632-258363	Sequence 258363, A
c 151	16	0.8	253	5	US-09-975-254-26610	Sequence 26610, A	c 224	16	0.8	587	6	US-10-027-632-363979	Sequence 363979, A
c 152	16	0.8	259	5	US-09-539-800C-2513	Sequence 2513, App	c 225	16	0.8	587	6	US-10-027-632-107850	Sequence 107850, A
c 153	16	0.8	264	5	US-09-975-254-29322	Sequence 29322, A	c 226	16	0.8	587	6	US-10-027-632-47971	Sequence 47971, A
c 154	16	0.8	266	6	US-10-027-632-270234	Sequence 270234, A	c 227	16	0.8	587	6	US-10-027-632-47972	Sequence 47972, A
c 155	16	0.8	270	6	US-10-099-926-1640	Sequence 1640, App	c 228	16	0.8	588	6	US-10-027-632-44003	Sequence 44003, A
c 156	16	0.8	274	5	US-09-539-331D-24963	Sequence 24963, A	c 229	16	0.8	588	6	US-10-027-632-44004	Sequence 44004, A
c 157	16	0.8	276	5	US-09-540-210B-26442	Sequence 26442, A	c 230	16	0.8	588	6	US-10-027-632-44005	Sequence 44005, A
c 158	16	0.8	287	6	US-10-099-926-1718	Sequence 1718, App	c 231	16	0.8	588	6	US-10-027-632-44006	Sequence 44006, A
c 159	16	0.8	292	6	US-10-027-632-65077	Sequence 65077, A	c 232	16	0.8	589	6	US-10-027-632-30096	Sequence 30096, A
c 160	16	0.8	309	5	US-09-899-575-93	Sequence 93, App1	c 233	16	0.8	589	6	US-10-027-632-42858	Sequence 42858, A
c 161	16	0.8	311	5	US-09-539-800C-9580	Sequence 9580, App	c 234	16	0.8	589	6	US-10-027-632-42859	Sequence 42859, A
c 162	16	0.8	314	5	US-09-721-544-16447	Sequence 16447, A	c 235	16	0.8	609	6	US-10-027-632-107012	Sequence 107012, A
c 163	16	0.8	315	5	US-09-721-544-20031	Sequence 20031, A	c 236	16	0.8	609	6	US-10-027-632-254312	Sequence 254312, A
c 164	16	0.8	335	6	US-10-027-632-53601	Sequence 53601, A	c 237	16	0.8	612	6	US-10-027-632-280430	Sequence 280430, A
c 165	16	0.8	337	5	US-09-721-544-14946	Sequence 14946, A	c 238	16	0.8	612	6	US-10-027-632-280431	Sequence 280431, A
c 166	16	0.8	353	5	US-09-721-544-10664	Sequence 10664, A	c 239	16	0.8	618	6	US-10-027-632-101574	Sequence 101574, A
c 167	16	0.8	354	5	US-09-721-544-13294	Sequence 13294, A	c 240	16	0.8	618	6	US-10-027-632-101575	Sequence 101575, A
c 168	16	0.8	364	5	US-09-721-544-17222	Sequence 17222, A	c 241	16	0.8	623	6	US-10-027-632-197807	Sequence 197807, A
c 169	16	0.8	367	1	PCT-US02-12612-186	Sequence 186, App	c 242	16	0.8	623	7	US-60-377-240-3511	Sequence 3511, App
c 170	16	0.8	367	6	US-10-125-968-186	Sequence 186, App	c 243	16	0.8	625	6	US-10-027-632-84363	Sequence 84363, A
c 171	16	0.8	373	5	US-09-721-544-18631	Sequence 18631, A	c 244	16	0.8	625	6	US-10-027-632-269318	Sequence 269318, A
c 172	16	0.8	386	1	PCT-US02-12612-905	Sequence 905, App	c 245	16	0.8	625	6	US-10-027-632-269319	Sequence 269319, A

C 246	16	0.8	626	6	US-10-027-632-78176	
C 247	16	0.8	630	6	US-10-027-632-27921	
C 248	16	0.8	630	6	US-10-027-632-308413	
C 249	16	0.8	637	6	US-10-027-632-218502	
C 250	16	0.8	637	6	US-10-027-632-228230	
C 251	16	0.8	639	6	US-10-027-632-285397	
C 252	16	0.8	643	6	US-10-027-632-267586	
C 253	16	0.8	650	6	US-10-027-632-188190	
C 254	16	0.8	650	6	US-10-027-632-231937	
C 255	16	0.8	650	6	US-10-027-632-126527	
C 256	16	0.8	651	6	US-10-027-632-126528	
C 257	16	0.8	654	6	US-10-027-632-44023	
C 258	16	0.8	657	6	US-10-027-632-101477	
C 259	16	0.8	659	6	US-10-027-632-228811	
C 260	16	0.8	659	6	US-10-027-632-228812	
C 261	16	0.8	660	6	US-10-027-632-2075	
C 262	16	0.8	660	6	US-10-027-632-228470	
C 263	16	0.8	661	6	US-10-027-632-248483	
C 264	16	0.8	664	6	US-10-027-632-204053	
C 265	16	0.8	664	6	US-10-027-632-204054	
C 266	16	0.8	668	6	US-10-027-632-7191	
C 267	16	0.8	668	6	US-10-027-632-224984	
C 268	16	0.8	668	6	US-10-027-632-228470	
C 269	16	0.8	668	6	US-10-027-632-228471	
C 270	16	0.8	668	6	US-10-027-632-228472	
C 271	16	0.8	668	6	US-10-027-632-228473	
C 272	16	0.8	668	6	US-10-027-632-228474	
C 273	16	0.8	668	6	US-10-027-632-228475	
C 274	16	0.8	668	6	US-10-027-632-151638	
C 275	16	0.8	668	6	US-10-027-632-151639	
C 276	16	0.8	668	6	US-10-027-632-151639	
C 277	16	0.8	668	6	US-10-027-632-151640	
C 278	16	0.8	697	6	US-10-027-632-32430	
C 279	16	0.8	713	6	US-10-027-632-25909	
C 280	16	0.8	718	6	US-10-027-632-164051	
C 281	16	0.8	718	6	US-10-027-632-164052	
C 282	16	0.8	722	6	US-10-027-632-151457	
C 283	16	0.8	725	5	US-09-634-754C-803	
C 284	16	0.8	732	6	US-10-027-632-33179	
C 285	16	0.8	737	6	US-10-027-632-28056	
C 286	16	0.8	737	6	US-10-027-632-28057	
C 287	16	0.8	737	6	US-10-027-632-98623	
C 288	16	0.8	739	6	US-10-027-632-101652	
C 289	16	0.8	760	6	US-10-027-632-119220	
C 290	16	0.8	773	6	US-10-027-632-29955	
C 291	16	0.8	774	6	US-10-027-632-12339	
C 292	16	0.8	780	6	US-10-027-632-162813	
C 293	16	0.8	785	6	US-10-027-632-174200	
C 294	16	0.8	785	6	US-10-027-632-174201	
C 295	16	0.8	789	6	US-10-027-632-156257	
C 300	16	0.8	801	6	US-10-027-632-7134	
C 301	16	0.8	816	6	US-10-027-632-157038	
C 302	16	0.8	816	6	US-10-027-632-170389	
C 303	16	0.8	824	6	US-10-027-632-157911	
C 304	16	0.8	824	6	US-10-027-632-157912	
C 305	16	0.8	826	6	US-10-027-632-134116	
C 306	16	0.8	856	6	US-10-027-632-120426	
C 307	16	0.8	856	6	US-10-027-632-120427	
C 308	16	0.8	863	6	US-10-027-632-167879	
C 309	16	0.8	863	6	US-10-027-632-167880	
C 310	16	0.8	866	6	US-10-027-632-120912	
C 311	16	0.8	873	7	US-60-360-039-46819	
C 312	16	0.8	884	7	US-60-360-039-37207	
C 313	16	0.8	884	5	US-09-919-002-4009	
C 314	16	0.8	907	6	US-10-027-632-153864	
C 315	16	0.8	907	6	US-10-027-632-153865	
C 316	16	0.8	907	6	US-10-027-632-153866	
C 317	16	0.8	947	5	US-09-634-754C-110	
C 318	16	0.8	970	6	US-10-027-632-100933	

Sequence 78176, A	319	16	0.8	991	6	US-10-027-632-252940	Sequence 252940,
Sequence 27921, A	320	16	0.8	994	6	US-10-027-632-31531	Sequence 31531, A
Sequence 308413,	C 321	16	0.8	1008	6	US-10-027-632-120200	Sequence 120200,
Sequence 218502,	C 322	16	0.8	1008	6	US-10-027-632-120201	Sequence 120201,
Sequence 229330,	C 323	16	0.8	1052	6	US-10-027-632-9749	Sequence 9749, Ap
Sequence 285937,	C 324	16	0.8	1052	6	US-10-027-632-9750	Sequence 9750, Ap
Sequence 267586,	C 325	16	0.8	1055	5	US-09-935-625-598	Sequence 598, App
Sequence 188190,	C 326	16	0.8	1055	5	US-09-935-625-4320	Sequence 4320, Ap
Sequence 231937,	C 327	16	0.8	1055	5	US-09-935-625-16506	Sequence 16506, A
Sequence 126527,	C 328	16	0.8	1059	6	US-10-027-632-30791	Sequence 30791, A
Sequence 126528,	C 329	16	0.8	1059	6	US-10-027-632-30792	Sequence 30792, A
Sequence 44023, A	C 330	16	0.8	1070	6	US-10-116-949-7	Sequence 7, App1
Sequence 101477,	C 331	16	0.8	1095	7	US-60-360-039-26628	Sequence 26628, A
Sequence 228811, A	C 332	16	0.8	1120	6	US-10-027-632-263041	Sequence 263041,
Sequence 228812, A	C 333	16	0.8	1120	6	US-10-027-632-263042	Sequence 263042,
Sequence 2077, Ap	C 334	16	0.8	1140	7	US-60-360-039-42145	Sequence 42145, A
Sequence 2075, Ap	C 335	16	0.8	1140	7	US-60-360-039-42145	Sequence 42145, A
Sequence 248483,	C 336	16	0.8	1161	7	US-60-360-039-43390	Sequence 43390, A
Sequence 204053,	C 337	16	0.8	1198	6	US-10-027-632-296806	Sequence 296806,
Sequence 204054,	C 338	16	0.8	1215	7	US-60-360-039-39282	Sequence 39282, A
Sequence 7191, Ap	C 339	16	0.8	1215	7	US-60-360-039-39659	Sequence 39659, A
Sequence 224984,	C 340	16	0.8	1215	7	US-60-360-039-40019	Sequence 40019, A
Sequence 228470,	C 341	16	0.8	1245	6	US-10-027-632-122677	Sequence 122677, A
Sequence 228471,	C 342	16	0.8	1245	6	US-10-027-632-122678	Sequence 122678,
Sequence 228472,	C 343	16	0.8	1245	6	US-10-027-632-122679	Sequence 122679,
Sequence 228473,	C 344	16	0.8	1275	6	US-10-027-632-109309	Sequence 109309,
Sequence 228474,	C 345	16	0.8	1275	6	US-10-027-632-208709	Sequence 208709,
Sequence 228475,	C 346	16	0.8	1292	5	US-09-935-625-9839	Sequence 9839, Ap
Sequence 151638,	C 347	16	0.8	1292	5	US-09-935-625-22159	Sequence 22159, A
Sequence 151639,	C 348	16	0.8	1292	5	US-09-935-625-22159	Sequence 22159, A
Sequence 151640,	C 349	16	0.8	1308	5	US-09-935-625-9833	Sequence 9833, Ap
Sequence 151457,	C 350	16	0.8	1308	5	US-09-935-625-22163	Sequence 22163, A
Sequence 324030, A	C 351	16	0.8	1308	7	US-60-360-039-26389	Sequence 26389, A
Sequence 25909, A	C 352	16	0.8	1314	5	US-09-935-625-8607	Sequence 8607, Ap
Sequence 164051,	C 353	16	0.8	1353	7	US-60-360-039-43674	Sequence 43674, Ap
Sequence 164052,	C 354	16	0.8	1365	7	US-60-360-039-47128	Sequence 47128, A
Sequence 151457,	C 355	16	0.8	1371	7	US-60-360-039-25373	Sequence 25373, A
Sequence 803, App	C 356	16	0.8	1443	7	US-60-360-039-35658	Sequence 35658, A
Sequence 33179, A	C 357	16	0.8	1537	6	US-10-144-781-43	Sequence 43, App1
Sequence 28056, A	C 358	16	0.8	1538	5	US-09-935-625-7727	Sequence 7727, Ap
Sequence 28057, A	C 359	16	0.8	1570	5	US-09-919-002-3621	Sequence 3621, Ap
Sequence 98623, A	C 360	16	0.8	1584	5	US-09-935-625-11555	Sequence 11555, A
Sequence 101652,	C 361	16	0.8	1584	5	US-09-935-625-19457	Sequence 19457, A
Sequence 119220,	C 362	16	0.8	1584	5	US-09-935-625-27965	Sequence 27965, A
Sequence 29255, A	C 363	16	0.8	1595	6	US-10-105-299-2514	Sequence 2514, Ap
Sequence 12339, A	C 364	16	0.8	1604	6	US-10-027-632-249488	Sequence 249488,
Sequence 162813,	C 365	16	0.8	1605	7	US-60-360-039-35603	Sequence 35603, A
Sequence 174200,	C 366	16	0.8	1632	7	US-60-360-039-1602	Sequence 1602, Ap
Sequence 174201,	C 367	16	0.8	1646	1	PCT-US02-09188-1602	Sequence 1692, Ap
Sequence 156257,	C 368	16	0.8	1646	6	US-10-105-2299-7235	Sequence 7235, Ap
Sequence 7134, Ap	C 369	16	0.8	1652	1	PCT-US02-09188-335	Sequence 335, App
Sequence 156402,	C 370	16	0.8	1652	1	PCT-US02-09370-352	Sequence 352, App
Sequence 156403,	C 371	16	0.8	1652	6	US-10-105-299-466	Sequence 466, App
Sequence 156404,	C 372	16	0.8	1656	7	US-60-360-039-36017	Sequence 36017, A
Sequence 170387,	C 373	16	0.8	1656	7	US-60-360-039-98305	Sequence 98305, A
Sequence 170388,	C 374	16	0.8	1659	6	US-10-027-632-41349	Sequence 41349, A
Sequence 170389,	C 375	16	0.8	1698	7	US-60-360-039-41349	Sequence 41349, A
Sequence 157911,	C 376	16	0.8	1705	6	US-09-991-150-506	Sequence 506, App
Sequence 157912,	C 377	16	0.8	1705	6	US-10-121-062-393	Sequence 393, App
Sequence 134116,	C 378	16	0.8	1709	6	US-10-104-047-51	Sequence 51, App1
Sequence 120426,	C 379	16	0.8	1727	7	US-60-360-039-46005	Sequence 46005, A
Sequence 120427,	C 380	16	0.8	1743	6	US-10-006-063A-103	Sequence 103, App
Sequence 167879,	C 381	16	0.8	1743	6	US-10-006-117A-103	Sequence 103, App
Sequence 167880,	C 382	16	0.8	1743	6	US-10-006-117A-103	Sequence 103, App
Sequence 120912,	C 383	16	0.8	1743	6	US-10-006-172A-103	Sequence 103, App
Sequence 46819, A	C 384	16	0.8	1743	6	US-10-006-768A-103	Sequence 103, App
Sequence 37207, A	C 385	16	0.8	1743	6	US-10-017-522A-103	Sequence 103, App
Sequence 4009, Ap	C 386	16	0.8	1743	6	US-10-017-610A-103	Sequence 103, App
Sequence 153864,	C 387	16	0.8	1743	6	US-10-006-041A-103	Sequence 103, App
Sequence 153865,	C 388	16	0.8	1743	6	US-10-006-818A-103	Sequence 103, App
Sequence 153866,	C 389	16	0.8	1743	6	US-10-012-121A-103	Sequence 103, App
Sequence 110, App	C 390	16	0.8	1743	6	US-10-015-386A-103	Sequence 103, App
Sequence 100933,	C 391	16	0.8	1743	6	US-10-015-387A-103	Sequence 103, App

538	16	0.8	1743	6	US-10-063-723-75	Sequence 75, Appl	611	16	0.8	5320	5	US-09-442-384B-725	Sequence 725, App
539	16	0.8	1743	6	US-10-063-724-75	Sequence 75, Appl	612	16	0.8	6882	6	US-10-105-299-8509	Sequence 8509, Ap
540	16	0.8	1743	6	US-10-063-726-75	Sequence 75, Appl	613	16	0.8	8297	6	US-10-105-299-7938	Sequence 7938, Ap
541	16	0.8	1743	6	US-10-063-727-75	Sequence 75, Appl	614	16	0.8	8297	6	US-10-105-299-14263	Sequence 14263, A
542	16	0.8	1743	6	US-10-063-728-75	Sequence 75, Appl	615	16	0.8	10835	6	US-10-105-299-14988	Sequence 14988, A
543	16	0.8	1743	6	US-10-063-730-75	Sequence 75, Appl	616	16	0.8	10912	6	US-10-144-781-13	Sequence 13, Appl
544	16	0.8	1743	6	US-10-063-731-75	Sequence 75, Appl	617	16	0.8	11532	6	US-10-105-299-12161	Sequence 12161, A
545	16	0.8	1743	6	US-10-063-733-75	Sequence 75, Appl	618	16	0.8	11532	6	US-10-116-355-753	Sequence 753, App
546	16	0.8	1743	6	US-10-063-734-75	Sequence 75, Appl	619	16	0.8	13790	5	US-09-919-002-13473	Sequence 12473, A
547	16	0.8	1743	6	US-10-063-735-75	Sequence 75, Appl	620	16	0.8	14333	6	US-10-105-299-9643	Sequence 9643, Ap
548	16	0.8	1743	6	US-10-063-736-75	Sequence 75, Appl	621	16	0.8	14333	6	US-10-105-299-14987	Sequence 14987, A
549	16	0.8	1743	6	US-10-063-741-75	Sequence 75, Appl	622	16	0.8	15212	6	US-10-105-299-9506	Sequence 9506, Ap
550	16	0.8	1743	6	US-10-063-742-75	Sequence 75, Appl	623	16	0.8	17861	6	US-10-105-299-9930	Sequence 9930, Ap
551	16	0.8	1743	6	US-10-063-744-75	Sequence 75, Appl	624	16	0.8	21799	6	US-10-105-299-10314	Sequence 10314, A
552	16	0.8	1743	6	US-10-063-747-75	Sequence 75, Appl	625	16	0.8	21799	6	US-10-116-355-662	Sequence 662, App
553	16	0.8	1743	6	US-10-063-745-75	Sequence 75, Appl	626	16	0.8	25307	5	US-09-919-002-12002	Sequence 12002, A
554	16	0.8	1743	6	US-10-013-907A-103	Sequence 103, App	627	16	0.8	34555	6	US-10-117-982-479	Sequence 479, App
555	16	0.8	1743	6	US-10-015-391A-103	Sequence 103, App	628	16	0.8	35408	6	US-10-155-649-3	Sequence 3, Appl
556	16	0.8	1743	6	US-10-015-499A-103	Sequence 103, App	629	16	0.8	36620	1	PCT-US01-28861-35	Sequence 30, Appl
557	16	0.8	1743	6	US-10-017-867A-103	Sequence 103, App	630	16	0.8	37474	1	PCT-US01-28861-25	Sequence 25, Appl
558	16	0.8	1743	6	US-10-006-856A-103	Sequence 103, App	631	16	0.8	38519	1	PCT-US01-28861-38	Sequence 28, Appl
559	16	0.8	1757	1	PCT-US02-12851-2	Sequence 2, Appl	632	16	0.8	52156	5	US-09-578-519B-18	Sequence 18, Appl
560	16	0.8	1757	6	US-10-116-993-2	Sequence 2, Appl	633	16	0.8	160820	1	US-09-578-519B-18	Sequence 5, Appl
561	16	0.8	1758	6	US-10-144-850-32	Sequence 32, Appl	634	16	0.8	536165	5	US-09-939-964A-1	Sequence 1, Appl
562	16	0.8	1761	7	US-60-360-039-25611	Sequence 25611, A	635	16	0.8	1191139	5	US-10-140-924-1	Sequence 1, Appl
563	16	0.8	1793	5	US-09-919-002-12532	Sequence 12532, A	636	16	0.8	1191139	6	US-10-027-632-179264	Sequence 179264, A
564	16	0.8	1805	6	US-10-105-299-14869	Sequence 14869, A	637	16	0.8	1223197	6	US-10-067-514-1	Sequence 1, Appl
565	16	0.8	1836	7	US-60-360-039-25634	Sequence 25634, A	638	16	0.8	1691139	6	US-10-067-514-1	Sequence 1, Appl
566	16	0.8	1842	7	US-60-360-039-37001	Sequence 37001, A	639	16	0.8	1691139	6	US-10-067-514-1	Sequence 1, Appl
567	16	0.8	1844	6	US-10-027-633-97948	Sequence 97948, A	640	16	0.8	1691139	6	US-10-067-514-1	Sequence 1, Appl
568	16	0.8	1944	6	US-10-027-633-97949	Sequence 97949, A	641	15	0.8	27	6	US-10-027-632-174961	Sequence 174961, A
569	16	0.8	1944	6	US-10-027-633-98175	Sequence 98175, A	642	15	0.8	88	5	US-09-539-331D-38821	Sequence 38821, A
570	16	0.8	1944	6	US-10-027-633-98176	Sequence 98176, A	643	15	0.8	97	5	US-09-539-331D-38821	Sequence 38821, A
571	16	0.8	1944	6	US-10-027-633-98177	Sequence 98177, A	644	15	0.8	109	5	US-09-539-331D-38821	Sequence 38821, A
572	16	0.8	1990	5	US-09-919-002-6691	Sequence 6691, Ap	645	15	0.8	136	5	US-09-540-210B-10017	Sequence 10017, A
573	16	0.8	2161	6	US-10-102-806-175	Sequence 175, App	646	15	0.8	155	5	US-09-539-331D-31261	Sequence 31261, A
574	16	0.8	2184	6	US-10-027-633-97571	Sequence 97571, A	647	15	0.8	169	6	US-10-141-324-1073	Sequence 1073, Ap
575	16	0.8	2188	6	US-10-099-926-1716	Sequence 1716, Ap	648	15	0.8	172	5	US-09-539-331D-8794	Sequence 8794, Ap
576	16	0.8	2197	6	US-10-116-355-606	Sequence 606, App	649	15	0.8	184	5	US-09-539-331D-36413	Sequence 36413, A
577	16	0.8	2269	5	US-10-099-926-1693	Sequence 1693, Ap	650	15	0.8	193	5	US-09-540-210B-22307	Sequence 22307, A
578	16	0.8	2269	5	US-10-099-926-1693	Sequence 1693, Ap	651	15	0.8	193	5	US-09-539-331D-28660	Sequence 28660, A
579	16	0.8	2271	5	US-09-935-625-21242	Sequence 935, App	652	15	0.8	196	5	US-09-540-210B-42336	Sequence 42336, A
580	16	0.8	2329	5	US-09-882-227-619	Sequence 619, App	653	15	0.8	199	5	US-09-539-331D-12296	Sequence 12296, A
581	16	0.8	2352	6	US-10-027-633-262544	Sequence 262544, A	654	15	0.8	205	5	US-09-540-210B-23470	Sequence 23470, A
582	16	0.8	2391	7	US-60-360-039-24390	Sequence 24390, A	655	15	0.8	207	5	US-09-540-209B-16085	Sequence 16085, A
583	16	0.8	2423	6	US-10-104-047-1378	Sequence 1378, Ap	656	15	0.8	207	5	US-09-540-210B-16232	Sequence 16232, A
584	16	0.8	2479	6	US-10-027-633-103567	Sequence 103567, A	657	15	0.8	208	5	US-09-539-331D-165125	Sequence 165125, A
585	16	0.8	2485	6	US-10-027-633-111968	Sequence 111968, A	658	15	0.8	208	5	US-09-539-331D-36125	Sequence 36125, A
586	16	0.8	2498	6	US-10-027-633-102064	Sequence 102064, A	659	15	0.8	210	5	US-09-540-210B-30086	Sequence 30086, A
587	16	0.8	2498	6	US-10-027-633-102065	Sequence 102065, A	660	15	0.8	213	5	US-09-540-210B-14593	Sequence 14593, A
588	16	0.8	2498	6	US-10-027-633-102066	Sequence 102066, A	661	15	0.8	215	5	US-09-540-210B-1593	Sequence 1593, A
589	16	0.8	2530	1	PCT-US02-13644-5	Sequence 5, Appl	662	15	0.8	216	5	US-09-899-575-81	Sequence 81, Appl
590	16	0.8	2530	6	US-10-136-819-5	Sequence 5, Appl	663	15	0.8	217	5	US-09-540-210B-11350	Sequence 11350, A
591	16	0.8	2532	6	US-10-104-047-1069	Sequence 1069, Ap	664	15	0.8	218	5	US-09-540-210B-14969	Sequence 14969, A
592	16	0.8	2570	6	US-10-102-524-1764	Sequence 1764, Ap	665	15	0.8	218	6	US-10-146-502-1073	Sequence 1073, Ap
593	16	0.8	2760	5	US-09-935-625-16899	Sequence 16899, A	666	15	0.8	220	5	US-09-540-210B-8893	Sequence 8893, Ap
594	16	0.8	2918	5	US-09-919-002-3785	Sequence 3785, Ap	667	15	0.8	220	5	US-09-540-210B-30219	Sequence 30219, A
595	16	0.8	2962	5	US-09-919-002-1374	Sequence 1374, Ap	668	15	0.8	221	6	US-09-142-569-319	Sequence 319, App
596	16	0.8	3051	6	US-10-027-633-256743	Sequence 256743, A	669	15	0.8	226	5	US-09-973-234-17028	Sequence 17028, A
597	16	0.8	3051	6	US-10-027-633-256744	Sequence 256744, A	670	15	0.8	226	5	US-09-460-592B-1025	Sequence 1025, Ap
598	16	0.8	3183	7	US-60-360-039-35634	Sequence 35634, A	671	15	0.8	228	5	US-09-539-331D-4571	Sequence 4571, Ap
599	16	0.8	3416	5	US-09-930-213-293	Sequence 293, App	672	15	0.8	228	5	US-09-539-331D-5754	Sequence 5754, App
600	16	0.8	3497	6	US-10-027-633-114771	Sequence 114771, A	673	15	0.8	232	5	US-09-539-806B-755	Sequence 755, App
601	16	0.8	3608	6	US-10-121-062-433	Sequence 433, App	674	15	0.8	232	5	US-09-539-800C-3477	Sequence 3477, Ap
602	16	0.8	3692	6	US-10-106-698-824	Sequence 824, App	675	15	0.8	232	5	US-09-539-800C-8906	Sequence 8906, Ap
603	16	0.8	3754	6	US-10-104-047-530	Sequence 530, App	676	15	0.8	233	5	US-09-975-254-19113	Sequence 19113, A
604	16	0.8	3777	6	US-09-919-002-3576	Sequence 3576, Ap	677	15	0.8	233	5	US-09-539-331D-16546	Sequence 16546, A
605	16	0.8	4022	6	US-10-104-047-293	Sequence 293, App	678	15	0.8	233	5	US-09-539-331D-40504	Sequence 40504, A
606	16	0.8	4054	6	US-10-011-585A-32	Sequence 32, Appl	679	15	0.8	234	5	US-09-539-331D-17931	Sequence 17931, A
607	16	0.8	4145	5	US-09-053-375B-628	Sequence 628, App	680	15	0.8	235	5	US-09-539-331D-25965	Sequence 25965, A
608	16	0.8	4150	6	US-10-104-047-963	Sequence 963, App	681	15	0.8	236	5	US-09-539-800C-3992	Sequence 3992, Ap
609	16	0.8	4695	5	US-09-919-002-7377	Sequence 7377, Ap	682	15	0.8	236	6	US-10-027-633-289083	Sequence 289083, A
610	16	0.8	5188	6	US-10-105-299-9507	Sequence 9507, Ap	683	15	0.8	236	6	US-10-027-633-289084	Sequence 289084, A

c 684	15	0.8	241	5	US-09-540-2108-26184	Sequence 26184, A	c 757	15	0.8	351	5	US-09-895-913A-279	Sequence 279, App
c 685	15	0.8	242	5	US-09-975-254-24704	Sequence 24704, A	c 758	15	0.8	355	5	US-09-721-544-17650	Sequence 17650, A
c 686	15	0.8	243	5	US-09-975-254-4356	Sequence 4356, Ap	c 759	15	0.8	356	6	US-10-102-524-10408	Sequence 10408, Ap
c 687	15	0.8	243	5	US-09-539-331D-23443	Sequence 23443, A	c 760	15	0.8	359	6	US-10-103-313-149	Sequence 149, App
c 688	15	0.8	244	5	US-09-975-254-17067	Sequence 17067, A	c 761	15	0.8	359	6	US-10-027-632-269643	Sequence 269643, A
c 689	15	0.8	244	5	US-09-975-254-30774	Sequence 30774, A	c 762	15	0.8	359	6	US-10-027-632-269644	Sequence 269644, A
c 690	15	0.8	245	5	US-09-539-800C-18974	Sequence 18974, A	c 763	15	0.8	359	6	US-10-027-632-269645	Sequence 269645, A
c 691	15	0.8	246	5	US-09-539-331D-30196	Sequence 30196, A	c 764	15	0.8	360	5	US-09-721-544-16530	Sequence 16530, A
c 692	15	0.8	247	5	US-09-975-254-7622	Sequence 7622, Ap	c 765	15	0.8	368	5	US-09-721-544-8351	Sequence 8351, Ap
c 693	15	0.8	248	5	US-09-975-254-11792	Sequence 11792, A	c 766	15	0.8	368	5	US-09-721-544-13779	Sequence 13779, A
c 694	15	0.8	249	5	US-09-540-2108-1970	Sequence 1970, Ap	c 767	15	0.8	368	5	US-09-919-002-7642	Sequence 7642, Ap
c 695	15	0.8	250	5	US-09-540-2108-4520	Sequence 4520, Ap	c 768	15	0.8	369	7	US-60-360-039-37174	Sequence 37174, A
c 696	15	0.8	251	5	US-09-539-331D-8680	Sequence 8680, Ap	c 769	15	0.8	370	5	US-09-721-544-10057	Sequence 10057, A
c 697	15	0.8	251	5	US-09-539-800C-2895	Sequence 2895, Ap	c 770	15	0.8	374	5	US-09-539-331D-6828	Sequence 6828, Ap
c 698	15	0.8	252	5	US-09-975-254-13665	Sequence 13665, A	c 771	15	0.8	379	5	US-09-991-936-1111	Sequence 1111, Ap
c 699	15	0.8	253	5	US-09-539-331D-19235	Sequence 19235, A	c 772	15	0.8	380	5	US-09-919-002-10515	Sequence 10515, A
c 700	15	0.8	255	5	US-09-975-254-19783	Sequence 19783, A	c 773	15	0.8	387	5	US-09-919-002-11228	Sequence 11228, A
c 701	15	0.8	256	5	US-09-975-254-5198	Sequence 5198, Ap	c 774	15	0.8	390	5	US-09-721-544-23477	Sequence 23477, A
c 702	15	0.8	256	5	US-09-975-254-12792	Sequence 12792, A	c 775	15	0.8	392	5	US-09-721-544-14570	Sequence 14570, A
c 703	15	0.8	257	5	US-09-540-2108-7130	Sequence 7130, Ap	c 776	15	0.8	393	5	US-09-540-209B-4763	Sequence 4763, Ap
c 704	15	0.8	258	5	US-09-975-254-16817	Sequence 16817, A	c 777	15	0.8	393	5	US-09-919-002-10269	Sequence 10269, A
c 705	15	0.8	258	5	US-09-975-254-19449	Sequence 19449, A	c 778	15	0.8	394	5	US-09-171-937C-47	Sequence 47, Ap
c 706	15	0.8	260	5	US-09-975-254-24904	Sequence 24904, A	c 779	15	0.8	395	5	US-09-689-809-699	Sequence 699, App
c 707	15	0.8	260	5	US-09-941-492-73	Sequence 73, Ap	c 780	15	0.8	395	5	US-09-919-002-10189	Sequence 10189, A
c 708	15	0.8	262	5	US-09-539-331D-31245	Sequence 31245, A	c 781	15	0.8	396	6	US-10-027-632-183302	Sequence 183302, A
c 709	15	0.8	263	5	US-09-975-254-26125	Sequence 26125, A	c 782	15	0.8	399	6	US-10-114-893-320	Sequence 320, App
c 710	15	0.8	263	5	US-09-539-331D-7671	Sequence 7671, Ap	c 783	15	0.8	401	7	US-60-317-240-9854	Sequence 9854, App
c 711	15	0.8	264	5	US-09-975-254-11024	Sequence 11024, A	c 784	15	0.8	403	5	US-09-539-800C-14291	Sequence 14291, A
c 712	15	0.8	264	5	US-09-975-254-11036	Sequence 11036, A	c 785	15	0.8	403	6	US-10-113-872-1201	Sequence 1201, Ap
c 713	15	0.8	265	5	US-09-540-2108-12835	Sequence 12835, A	c 786	15	0.8	411	1	PCT-US02-13378-136	Sequence 136, App
c 714	15	0.8	266	5	US-09-975-254-2294	Sequence 2294, Ap	c 787	15	0.8	411	1	US-10-124-805-136	Sequence 136, App
c 715	15	0.8	266	5	US-09-539-331D-14182	Sequence 14182, A	c 788	15	0.8	412	6	US-10-011-154-302	Sequence 302, App
c 716	15	0.8	266	5	US-09-539-800C-18576	Sequence 18576, A	c 789	15	0.8	414	6	US-10-007-926-55	Sequence 55, Ap
c 717	15	0.8	266	5	US-09-539-800C-18621	Sequence 18621, A	c 790	15	0.8	415	1	PCT-US02-10421-2068	Sequence 2068, Ap
c 718	15	0.8	267	5	US-09-975-254-6533	Sequence 6533, Ap	c 791	15	0.8	415	6	US-10-112-699-2068	Sequence 2068, Ap
c 719	15	0.8	267	5	US-09-540-209B-3348	Sequence 3348, Ap	c 792	15	0.8	415	6	US-10-027-632-183225	Sequence 183225, A
c 720	15	0.8	269	5	US-09-539-331D-295	Sequence 295, App	c 793	15	0.8	417	6	US-10-011-154-186	Sequence 186, App
c 721	15	0.8	269	5	US-09-539-806B-3157	Sequence 3157, Ap	c 794	15	0.8	418	6	US-10-027-632-299100	Sequence 299100, A
c 722	15	0.8	270	5	US-09-975-254-21506	Sequence 21506, A	c 795	15	0.8	419	5	US-09-919-002-5504	Sequence 5504, Ap
c 723	15	0.8	270	5	US-09-539-331D-38553	Sequence 38553, A	c 796	15	0.8	419	6	US-10-027-632-70184	Sequence 70184, A
c 724	15	0.8	271	5	US-09-539-331D-802	Sequence 802, App	c 797	15	0.8	420	6	US-10-027-632-277917	Sequence 277917, A
c 725	15	0.8	271	5	US-09-539-331D-13423	Sequence 13423, A	c 798	15	0.8	420	6	US-10-027-632-277918	Sequence 277918, A
c 726	15	0.8	272	5	US-09-975-254-9085	Sequence 9085, A	c 799	15	0.8	422	6	US-10-146-502-1828	Sequence 1828, Ap
c 727	15	0.8	277	5	US-09-539-800C-9689	Sequence 9689, Ap	c 800	15	0.8	423	6	US-10-027-632-43145	Sequence 43145, A
c 728	15	0.8	280	5	US-09-540-2108-19436	Sequence 19436, A	c 801	15	0.8	426	6	US-10-027-632-86524	Sequence 86524, A
c 729	15	0.8	280	5	US-09-539-331D-19436	Sequence 19436, A	c 802	15	0.8	426	6	US-10-027-632-86525	Sequence 86525, A
c 730	15	0.8	282	5	US-09-540-2108-13071	Sequence 13071, A	c 803	15	0.8	426	6	US-10-027-632-86526	Sequence 86526, A
c 731	15	0.8	286	5	US-09-539-331D-3325	Sequence 3325, Ap	c 804	15	0.8	426	6	US-10-027-632-316406	Sequence 316406, A
c 732	15	0.8	287	5	US-09-540-2108-33574	Sequence 33574, A	c 805	15	0.8	426	6	US-10-027-632-316407	Sequence 316407, A
c 733	15	0.8	287	5	US-09-539-806B-1493	Sequence 1493, Ap	c 806	15	0.8	426	6	US-10-027-632-316408	Sequence 316408, A
c 734	15	0.8	288	5	US-09-539-800C-13243	Sequence 13243, A	c 807	15	0.8	430	5	US-09-721-544-21591	Sequence 21591, A
c 735	15	0.8	292	1	PCT-US02-09188-1659	Sequence 1659, Ap	c 808	15	0.8	430	6	US-10-011-154-4557	Sequence 4557, Ap
c 736	15	0.8	292	1	PCT-US02-09370-1762	Sequence 1762, Ap	c 809	15	0.8	433	6	US-10-027-632-183144	Sequence 183144, A
c 737	15	0.8	292	5	US-09-540-2108-10133	Sequence 10133, A	c 810	15	0.8	433	6	US-10-027-632-183145	Sequence 183145, A
c 738	15	0.8	292	6	US-10-105-299-7325	Sequence 7325, Ap	c 811	15	0.8	433	6	US-10-027-632-183146	Sequence 183146, A
c 739	15	0.8	293	5	US-09-539-331D-31221	Sequence 31221, A	c 812	15	0.8	436	5	US-09-991-936-1630	Sequence 1630, Ap
c 740	15	0.8	294	5	US-09-540-2108-20446	Sequence 20446, A	c 813	15	0.8	436	6	US-10-011-154-4221	Sequence 4221, Ap
c 741	15	0.8	295	6	US-10-143-775-89	Sequence 89, Ap	c 814	15	0.8	437	5	US-09-721-544-21942	Sequence 21942, A
c 742	15	0.8	300	5	US-09-540-2108-35625	Sequence 35625, A	c 815	15	0.8	438	5	US-09-539-800C-16509	Sequence 16509, A
c 743	15	0.8	300	5	US-09-539-331D-35308	Sequence 35308, A	c 816	15	0.8	440	7	US-60-360-039-30766	Sequence 30766, A
c 744	15	0.8	301	6	US-10-027-632-35966	Sequence 35966, A	c 817	15	0.8	443	5	US-09-721-544-21331	Sequence 21331, A
c 745	15	0.8	301	6	US-10-027-632-35966	Sequence 35966, A	c 818	15	0.8	443	5	US-09-721-544-21331	Sequence 21331, A
c 746	15	0.8	308	6	US-10-144-654-6	Sequence 65285, A	c 819	15	0.8	443	6	US-10-027-632-268517	Sequence 268517, A
c 747	15	0.8	309	5	US-09-899-575-94	Sequence 94, Ap	c 820	15	0.8	448	6	US-10-027-632-65820	Sequence 65820, A
c 748	15	0.8	310	5	US-09-539-331D-30380	Sequence 30380, A	c 821	15	0.8	448	6	US-10-027-632-65821	Sequence 65821, A
c 749	15	0.8	311	5	US-09-919-002-9159	Sequence 9159, Ap	c 822	15	0.8	448	6	US-10-027-632-66538	Sequence 66538, A
c 750	15	0.8	312	5	US-09-540-2108-10631	Sequence 10631, A	c 823	15	0.8	448	6	US-10-027-632-66539	Sequence 66539, A
c 751	15	0.8	318	6	US-10-116-355-644	Sequence 644, App	c 824	15	0.8	448	6	US-10-027-632-67466	Sequence 67466, A
c 752	15	0.8	323	5	US-09-941-492-102	Sequence 102, App	c 825	15	0.8	448	6	US-10-027-632-67467	Sequence 67467, A
c 753	15	0.8	335	5	US-09-539-331D-16731	Sequence 16731, A	c 826	15	0.8	448	6	US-10-027-632-67876	Sequence 67876, A
c 754	15	0.8	343	5	US-09-539-800C-9775	Sequence 9775, Ap	c 827	15	0.8	448	6	US-10-027-632-67877	Sequence 67877, A
c 755	15	0.8	347	5	US-09-721-544-20142	Sequence 20142, A	c 828	15	0.8	448	6	US-10-027-632-183658	Sequence 183658, A
c 756	15	0.8	348	5	US-09-721-544-10674	Sequence 10674, A	c 829	15	0.8	450	6	US-10-027-632-183692	Sequence 183692, A

C 830	15	0.8	450	6	US-10-027-632-78693	Sequence 78693, A	C 903	15	0.8	507	7	US-60-377-240-6237	Sequence 6237, Ap
C 831	15	0.8	450	6	US-10-027-632-314577	Sequence 314577, A	C 904	15	0.8	508	5	US-09-721-544-4927	Sequence 4927, Ap
C 832	15	0.8	450	6	US-10-027-632-314578	Sequence 314578, A	C 905	15	0.8	508	5	US-10-027-632-96659	Sequence 96659, A
C 833	15	0.8	452	6	US-10-011-154-1036	Sequence 1036, Ap	C 906	15	0.8	508	6	US-10-027-632-320980	Sequence 320980, A
C 834	15	0.8	453	5	US-09-620-393B-6564	Sequence 6564, Ap	C 907	15	0.8	510	6	US-10-027-632-78300	Sequence 78300, A
C 835	15	0.8	453	5	US-09-721-544-21933	Sequence 21933, A	C 908	15	0.8	510	6	US-10-027-632-78301	Sequence 78301, A
C 836	15	0.8	454	6	US-10-027-632-213709	Sequence 213709, A	C 909	15	0.8	510	6	US-10-027-632-78302	Sequence 78302, A
C 837	15	0.8	455	6	US-10-027-632-35196	Sequence 35196, A	C 910	15	0.8	510	6	US-10-027-632-79199	Sequence 79199, A
C 838	15	0.8	455	6	US-10-027-632-195204	Sequence 195204, A	C 911	15	0.8	510	6	US-10-027-632-79200	Sequence 79200, A
C 839	15	0.8	455	6	US-10-027-632-209893	Sequence 209893, A	C 912	15	0.8	510	6	US-10-027-632-79201	Sequence 79201, A
C 840	15	0.8	457	5	US-09-732-560-1	Sequence 1, Appl1	C 913	15	0.8	510	6	US-10-027-632-87203	Sequence 87203, A
C 841	15	0.8	458	1	PCT-US01-44838-1209	Sequence 1209, Ap	C 914	15	0.8	510	6	US-10-027-632-87204	Sequence 87204, A
C 842	15	0.8	458	6	US-10-027-632-288734	Sequence 288734, A	C 915	15	0.8	510	6	US-10-027-632-300742	Sequence 300742, A
C 843	15	0.8	459	5	US-09-919-002-11340	Sequence 11340, A	C 916	15	0.8	510	6	US-10-027-632-300743	Sequence 300743, A
C 844	15	0.8	459	6	US-10-027-632-85064	Sequence 85064, A	C 917	15	0.8	510	6	US-10-027-632-306912	Sequence 306912, A
C 845	15	0.8	460	5	US-09-919-002-6316	Sequence 6316, Ap	C 918	15	0.8	510	6	US-10-027-632-316606	Sequence 316606, A
C 846	15	0.8	460	6	US-10-027-632-81367	Sequence 81367, A	C 919	15	0.8	510	6	US-10-027-632-316607	Sequence 316607, A
C 847	15	0.8	462	6	US-10-027-632-209561	Sequence 209561, A	C 920	15	0.8	511	6	US-10-027-632-92073	Sequence 92073, A
C 848	15	0.8	466	7	US-60-377-240-7366	Sequence 7366, Ap	C 921	15	0.8	511	6	US-10-027-632-317886	Sequence 317886, A
C 849	15	0.8	468	5	US-09-689-909-1136	Sequence 1136, Ap	C 922	15	0.8	512	5	US-09-919-002-77209	Sequence 77209, Ap
C 850	15	0.8	468	6	US-10-027-632-194195	Sequence 194195, A	C 923	15	0.8	512	6	US-10-027-632-180687	Sequence 180687, A
C 851	15	0.8	468	6	US-10-027-632-194196	Sequence 194196, A	C 924	15	0.8	512	6	US-10-027-632-180688	Sequence 180688, A
C 852	15	0.8	470	6	US-10-027-632-28695	Sequence 28695, A	C 925	15	0.8	513	5	US-09-393B-7126	Sequence 7126, Ap
C 853	15	0.8	471	5	US-09-732-560-100	Sequence 100, Ap	C 926	15	0.8	514	5	US-09-539-800C-14552	Sequence 14552, A
C 854	15	0.8	471	5	US-09-721-544-11919	Sequence 11919, A	C 927	15	0.8	514	6	US-10-027-632-72540	Sequence 72540, A
C 855	15	0.8	473	6	US-10-027-632-273009	Sequence 273009, A	C 928	15	0.8	514	6	US-10-027-632-132275	Sequence 132275, A
C 856	15	0.8	474	6	US-10-027-632-70510	Sequence 70510, A	C 929	15	0.8	516	6	US-10-027-632-132276	Sequence 132276, A
C 857	15	0.8	476	6	US-10-027-632-181386	Sequence 181386, A	C 930	15	0.8	516	6	US-10-027-632-132277	Sequence 132277, A
C 858	15	0.8	476	6	US-10-027-632-181387	Sequence 181387, A	C 931	15	0.8	517	1	PCT-US02-10421-1219	Sequence 1219, Ap
C 859	15	0.8	476	6	US-10-027-632-181388	Sequence 181388, A	C 932	15	0.8	517	6	US-10-112-699-1219	Sequence 1219, Ap
C 860	15	0.8	477	6	US-10-027-632-34398	Sequence 34398, A	C 933	15	0.8	517	7	US-60-360-033-28798	Sequence 28798, A
C 861	15	0.8	477	6	US-10-027-632-34399	Sequence 34399, A	C 934	15	0.8	518	6	US-10-027-632-188217	Sequence 188217, A
C 862	15	0.8	479	6	US-10-027-632-64390	Sequence 64390, A	C 935	15	0.8	519	6	US-10-027-632-67043	Sequence 67043, A
C 863	15	0.8	480	5	US-09-721-544-11953	Sequence 11953, A	C 936	15	0.8	519	6	US-10-027-632-67044	Sequence 67044, A
C 864	15	0.8	481	5	US-09-721-544-7485	Sequence 7485, Ap	C 937	15	0.8	519	6	US-10-027-632-300803	Sequence 300803, A
C 865	15	0.8	482	7	US-60-377-240-5405	Sequence 5405, Ap	C 938	15	0.8	519	6	US-10-027-632-300804	Sequence 300804, A
C 866	15	0.8	483	5	US-09-539-331D-31802	Sequence 31802, A	C 939	15	0.8	520	5	US-09-620-393B-3374	Sequence 3374, A
C 867	15	0.8	483	5	US-09-721-544-7383	Sequence 7383, Ap	C 940	15	0.8	520	6	US-10-027-632-127843	Sequence 127843, A
C 868	15	0.8	483	6	US-10-027-632-83375	Sequence 83375, A	C 941	15	0.8	520	6	US-10-027-632-127844	Sequence 127844, A
C 869	15	0.8	485	5	US-09-721-544-3482	Sequence 3482, Ap	C 942	15	0.8	520	6	US-10-027-632-64867	Sequence 64867, A
C 870	15	0.8	485	6	US-10-027-632-42580	Sequence 42580, A	C 943	15	0.8	520	6	US-10-027-632-300594	Sequence 300594, A
C 871	15	0.8	489	5	US-09-919-002-458	Sequence 458, App	C 944	15	0.8	520	6	US-10-027-632-306895	Sequence 306895, A
C 872	15	0.8	489	7	US-60-377-240-4955	Sequence 4955, Ap	C 945	15	0.8	522	5	US-09-919-002-5108	Sequence 5108, Ap
C 873	15	0.8	493	6	US-10-027-632-274869	Sequence 274869, A	C 946	15	0.8	522	6	US-10-099-928-6659	Sequence 928, Ap
C 874	15	0.8	493	6	US-10-027-632-321124	Sequence 321124, A	C 947	15	0.8	522	6	US-10-027-632-96364	Sequence 96364, A
C 875	15	0.8	493	6	US-10-027-632-321126	Sequence 321126, A	C 948	15	0.8	522	6	US-10-027-632-96365	Sequence 96365, A
C 876	15	0.8	496	5	US-09-721-544-250	Sequence 250, App	C 949	15	0.8	522	6	US-10-027-632-127841	Sequence 127841, A
C 877	15	0.8	496	6	US-10-027-632-246996	Sequence 246996, A	C 950	15	0.8	522	6	US-10-027-632-127843	Sequence 127843, A
C 878	15	0.8	497	6	US-09-539-331D-35400	Sequence 35400, A	C 951	15	0.8	522	6	US-10-027-632-127844	Sequence 127844, A
C 879	15	0.8	497	6	US-10-027-632-35041	Sequence 35041, A	C 952	15	0.8	522	6	US-10-027-632-270671	Sequence 270671, A
C 880	15	0.8	497	6	US-10-027-632-301352	Sequence 301352, A	C 953	15	0.8	522	6	US-10-027-632-306195	Sequence 306195, A
C 881	15	0.8	498	1	PCT-US02-12612-670	Sequence 670, App	C 954	15	0.8	522	6	US-10-027-632-306196	Sequence 306196, A
C 882	15	0.8	498	6	US-09-991-936-187	Sequence 187, App	C 955	15	0.8	522	6	US-10-027-632-81742	Sequence 81742, A
C 883	15	0.8	498	6	US-10-125-968-670	Sequence 670, App	C 956	15	0.8	523	6	US-10-027-632-81743	Sequence 81743, A
C 884	15	0.8	499	7	US-60-360-033-29432	Sequence 29432, A	C 957	15	0.8	523	6	US-10-027-632-93575	Sequence 93575, A
C 885	15	0.8	500	5	US-09-991-936-1767	Sequence 1767, Ap	C 958	15	0.8	523	6	US-10-027-632-307795	Sequence 307795, A
C 886	15	0.8	500	5	US-09-919-002-4908	Sequence 4908, Ap	C 959	15	0.8	524	5	US-09-539-800C-8711	Sequence 8711, Ap
C 887	15	0.8	500	6	US-10-027-632-71118	Sequence 71118, A	C 960	15	0.8	524	6	US-10-027-632-3562	Sequence 3562, Ap
C 888	15	0.8	500	6	US-10-027-632-245162	Sequence 245162, A	C 961	15	0.8	525	7	US-60-377-240-1114	Sequence 1114, Ap
C 889	15	0.8	500	6	US-10-027-632-245163	Sequence 245163, A	C 962	15	0.8	526	6	US-10-027-632-40630	Sequence 40630, A
C 890	15	0.8	500	6	US-10-027-632-287952	Sequence 287952, A	C 963	15	0.8	526	6	US-10-027-632-40631	Sequence 40631, A
C 891	15	0.8	500	6	US-10-027-632-287953	Sequence 287953, A	C 964	15	0.8	526	6	US-10-027-632-40632	Sequence 40632, A
C 892	15	0.8	502	6	US-10-027-632-22320	Sequence 22320, A	C 965	15	0.8	529	5	US-09-539-331D-34371	Sequence 34371, A
C 893	15	0.8	502	6	US-10-027-632-22321	Sequence 22321, A	C 966	15	0.8	529	6	US-10-027-632-37512	Sequence 37512, A
C 894	15	0.8	503	6	US-10-027-632-6807	Sequence 6807, Ap	C 967	15	0.8	529	6	US-10-027-632-37513	Sequence 37513, A
C 895	15	0.8	504	6	US-10-027-632-43952	Sequence 43952, A	C 968	15	0.8	529	6	US-10-027-632-334266	Sequence 334266, A
C 896	15	0.8	504	6	US-10-027-632-75620	Sequence 75620, A	C 969	15	0.8	529	6	US-10-027-632-334267	Sequence 334267, A
C 897	15	0.8	505	5	US-09-721-544-4928	Sequence 4928, Ap	C 970	15	0.8	530	5	US-09-539-331D-35157	Sequence 35157, A
C 898	15	0.8	505	6	US-10-027-632-290728	Sequence 290728, A	C 971	15	0.8	531	6	US-10-027-632-124282	Sequence 124282, A
C 899	15	0.8	505	6	US-10-027-632-290729	Sequence 290729, A	C 972	15	0.8	531	6	US-10-027-632-134628	Sequence 134628, A
C 900	15	0.8	505	6	US-10-027-632-290730	Sequence 290730, A	C 973	15	0.8	534	5	US-09-919-002-5585	Sequence 5585, Ap
C 901	15	0.8	506	5	US-09-539-331D-35413	Sequence 35413, A	C 974	15	0.8	537	5	US-09-540-205B-3138	Sequence 3138, Ap
C 902	15	0.8	507	6	US-10-027-632-300634	Sequence 300634, A	C 975	15	0.8	537	6	US-10-027-632-4308	Sequence 4308, Ap

c 976	15	0.8	537	6	US-10-027-632-4909	Sequence 4909, Ap
977	15	0.8	537	7	US-60-360-039-42617	Sequence 42617, A
978	15	0.8	539	5	US-09-721-544-5625	Sequence 5625, Ap
979	15	0.8	540	6	US-10-027-632-54028	Sequence 54028, A
980	15	0.8	540	6	US-10-027-632-289703	Sequence 289703
981	15	0.8	540	6	US-10-027-632-314558	Sequence 314558
982	15	0.8	540	6	US-10-027-632-321910	Sequence 321910
983	15	0.8	541	6	US-10-027-632-26193	Sequence 26193, A
984	15	0.8	541	6	US-10-027-632-109123	Sequence 109123
985	15	0.8	541	6	US-10-027-632-321125	Sequence 321125
986	15	0.8	542	6	US-10-027-632-203476	Sequence 203476
987	15	0.8	542	6	US-10-027-632-287844	Sequence 287844
988	15	0.8	543	6	US-10-027-632-285403	Sequence 285403
989	15	0.8	544	5	US-09-721-544-5626	Sequence 5626, Ap
990	15	0.8	545	6	US-10-027-632-76960	Sequence 76960, A
991	15	0.8	545	6	US-10-027-632-185376	Sequence 185376
992	15	0.8	545	6	US-10-027-632-300241	Sequence 300241
993	15	0.8	545	6	US-10-144-850-189	Sequence 189, App
994	15	0.8	546	6	US-10-027-632-41529	Sequence 41529, A
995	15	0.8	546	6	US-10-027-632-66628	Sequence 66628, A
996	15	0.8	546	6	US-10-027-632-210228	Sequence 210228
997	15	0.8	546	6	US-10-027-632-247833	Sequence 247833
998	15	0.8	547	6	US-10-027-632-204285	Sequence 204285
999	15	0.8	547	6	US-10-027-632-204286	Sequence 204286
c1000	15	0.8	547	6	US-10-027-632-204287	Sequence 204287

ALIGNMENTS

RESULT 1
US-08-961-083-1
Sequence 1, Application US/08961083
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
TELECOMMUNICATION INFORMATION:
REFERENCE/DOCKET NUMBER: PB340P2
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1999 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-961-083-1

		Query Match	100.0%;	Score 1999;	DB 4;	Length 1999;
		Best Local Similarity	100.0%;	Pred. No. 0;		
		Matches 1999;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TAAATATGACATATAAATCACTATTCGTGATTTGGTTTCGAACGGCCGCCA	60			
Db	1	TAAATATGACATATAAATCACTATTCGTGATTTGGTTTCGAACGGCCGCCA	60			
Qy	61	TGCCAAGTAATGATTTCCACAGATTTGGTTAAAGCAATGTTCTATGCAAGACA	120			
Db	61	TGCCAAGTAATGATTTCCACAGATTTGGTTAAAGCAATGTTCTATGCAAGACA	120			
Qy	121	TGCGTTCTTGACACACAGGGGGATTGATACATCCGATTCCTGGAGCTTTTCGCCAA	180			
Db	121	TGCGTTCTTGACACACAGGGGGATTGATACATCCGATTCCTGGAGCTTTTCGCCAA	180			
Qy	181	TCGCAAGAATTTCCCTCCCAAGGTGATACCTCTACCCCAAGTGTATTAAGTTGAC	240			
Db	181	TCGCAAGAATTTCCCTCCCAAGGTGATACCTCTACCCCAAGTGTATTAAGTTGAC	240			
Qy	241	TTACTTTCAACTTCGACTCCGACAGACTATTTCTGTAAGGCTCAGAGAGCTTGGT	300			
Db	241	TTACTTTCAACTTCGACTCCGACAGACTATTTCTGTAAGGCTCAGAGAGCTTGGT	300			
Qy	301	AGCGATTCACTAGAACAAAGCAACCAAGCAAGAAATCTGACTACTATATAATTA	360			
Db	301	AGCGATTCACTAGAACAAAGCAACCAAGCAAGAAATCTGACTACTATATAATTA	360			
Qy	361	GGTCTACATGCTTAATGGAACCTATGATGACAGACAGCTCAAAACTACTATGTTAA	420			
Db	361	GGTCTACATGCTTAATGGAACCTATGATGACAGACAGCTCAAAACTACTATGTTAA	420			
Qy	421	AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTCTGCTGGTAATGCTCAGGACC	480			
Db	421	AGACCTCAATATTTAAGTTTACCTCAGTTAGCTTCTGCTGGTAATGCTCAGGACC	480			
Qy	481	AAACCAATATGACCCCTATTACATCCAGAGACGCCCAAGCCGCGAACTTGCTCT	540			
Db	481	AAACCAATATGACCCCTATTACATCCAGAGACGCCCAAGCCGCGAACTTGCTCT	540			
Qy	541	ATCTGAATGAAAAATCAAGGCTACATCTGCTGAACATATGAGAAACACTCAATAC	600			
Db	541	ATCTGAATGAAAAATCAAGGCTACATCTGCTGAACATATGAGAAACACTCAATAC	600			
Qy	601	ACCAATTACTGATGAGTACCAAGTCAATAGCAAGTAATTAACCTCTTACATGA	660			
Db	601	ACCAATTACTGATGAGTACCAAGTCAATAGCAAGTAATTAACCTCTTACATGA	660			
Qy	661	TAAATACCTCAAGGAAGTCAATCAATGTAAGAAAGAAAGAGCTATTAACCTACTAC	720			
Db	661	TAAATACCTCAAGGAAGTCAATCAATGTAAGAAAGAAAGAGCTATTAACCTACTAC	720			
Qy	721	AACCTGGATGATGCTACCAATGTAGACCAAGAGCTCAAAACATGTGGGGTAT	780			
Db	721	AACCTGGATGATGCTACCAATGTAGACCAAGAGCTCAAAACATGTGGGGTAT	780			
Qy	781	TTACAATACAGACGAATACCTGCTATCCAGAGATGAATTCGAAGTCCCTTACCAT	840			
Db	781	TTACAATACAGACGAATACCTGCTATCCAGAGATGAATTCGAAGTCCCTTACCAT	840			
Qy	841	TGTTGATGTTTCAACGGTAAGTCAATGGCCAGTGAAGACAGCCCATCAGTCAAGTAA	900			
Db	841	TGTTGATGTTTCAACGGTAAGTCAATGGCCAGTGAAGACAGCCCATCAGTCAAGTAA	900			
Qy	901	TGTTCTTGGGAATTAACCAAGCAGTAAGAAACAAACCGGAGTGGGATCAACTAGAA	960			
Db	901	TGTTCTTGGGAATTAACCAAGCAGTAAGAAACAAACCGGAGTGGGATCAACTAGAA	960			
Qy	961	ACCGATACAGACTATGCTCTGCTGGAGTACGGGTCTACAGATTCAACTGCTACTAT	1020			
Db	961	ACCGATACAGACTATGCTCTGCTGGAGTACGGGTCTACAGATTCAACTGCTACTAT	1020			
Qy	1021	CGTTCAAGATGAGCCCTATTAACCTGGGACAAATTAATCTGTTTATTAAGTGGATAG	1080			

```

|||||
Db 1021 CGTTACAGATGAGCCCTATTAACCTGAGGACAAATACCTGTTATTAACGGGATAG 1080
|||
Qy 1081 GGGCTACTTTGGCAACATCACTTGGCAATACGCCCTGGCAACAATCGCGAAGCTCCAGC 1140
|||
Db 1081 GGGCTACTTTGGCAACATCACTTGGCAATACGCCCTGGCAACAATCGCGAAGCTCCAGC 1140
|||
Qy 1141 CGTGAACACTCTTAACCAAGTGGGACTCAACCGCGCAAGACTTCTTAATGGTCTAGG 1200
|||
Db 1141 CGTGAACACTCTTAACCAAGTGGGACTCAACCGCGCAAGACTTCTTAATGGTCTAGG 1200
|||
Qy 1201 AATGCACTACCCAAAGTATTCATCTCAATGCAATGCAATTCGAATACACAACCGAATCAGA 1260
|||
Db 1201 AATGCACTACCCAAAGTATTCATCTCAATGCAATGCAATTCGAATACACAACCGAATCAGA 1260
|||
Qy 1261 CAAAAAATATGGAGCAAGTAGTAAAAAGATGGCTGCTTACGCTGCTTGGCAATG 1320
|||
Db 1261 CAAAAAATATGGAGCAAGTAGTAAAAAGATGGCTGCTTACGCTGCTTGGCAATG 1320
|||
Qy 1321 TGGAACTTACTATTAACCAATGATATCCATTAAGTGGTCTTATGATGGAGTGAAGA 1380
|||
Db 1321 TGGAACTTACTATTAACCAATGATATCCATTAAGTGGTCTTATGATGGAGTGAAGA 1380
|||
Qy 1381 AGAGTTCTCTAATGTGGAAGCTGTCATGAAGAAACGACAGCCTATATGATGACCGA 1440
|||
Db 1381 AGAGTTCTCTAATGTGGAAGCTGTCATGAAGAAACGACAGCCTATATGATGACCGA 1440
|||
Qy 1441 CATGATGAAAAACAGTCTTACTATGAACTGGACGAATGCCATCTTGGTCTCC 1500
|||
Db 1441 CATGATGAAAAACAGTCTTACTATGAACTGGACGAATGCCATCTTGGTCTCC 1500
|||
Qy 1501 TCAGGCTGGTAAAAACAGAACTCTAATATATACAGAGGAATGAAAAACCATCA 1560
|||
Db 1501 TCAGGCTGGTAAAAACAGAACTCTAATATATACAGAGGAATGAAAAACCATCA 1560
|||
Qy 1561 GACCTCTCAATTTGTAGACACCTGATGAACTATTTGCTGGCTATATACCGTAATATTCAT 1620
|||
Db 1561 GACCTCTCAATTTGTAGACACCTGATGAACTATTTGCTGGCTATATACCGTAATATTCAT 1620
|||
Qy 1621 GGGTATGAGACAGGCTATTTCAACCGTGTGACACCACTTGTGAAGAAAGCTTACGGT 1680
|||
Db 1621 GGGTATGAGACAGGCTATTTCAACCGTGTGACACCACTTGTGAAGAAAGCTTACGGT 1680
|||
Qy 1681 CGGTGCAAAAGTTTACCGCTCTATGATGACCTGCTGTGAAGAAAGCTTACCGT 1740
|||
Db 1681 CGGTGCAAAAGTTTACCGCTCTATGATGACCTGCTGTGAAGAAAGCTTACCGT 1740
|||
Qy 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGGTCTCG 1800
|||
Db 1741 TTGGAATATACAGAGGGGCTCTACAGAAATGAGAAATTCGTAATTTAAAAATGGTCTCG 1800
|||
Qy 1801 TTCTAGTGGAACTCAACGCTCTCAACAAACACCCCATCAACGTGAAGTCAAGCTCAGC 1860
|||
Db 1801 TTCTAGTGGAACTCAACGCTCTCAACAAACACCCCATCAACGTGAAGTCAAGCTCAGC 1860
|||
Qy 1861 ATCAGATAGTCAACTTCAAGTCTAGCTCAACCACTCCAAAGCAAAATATAGTACGAC 1920
|||
Db 1861 ATCAGATAGTCAACTTCAAGTCTAGCTCAACCACTCCAAAGCAAAATATAGTACGAC 1920
|||
Qy 1921 TACCAATCTCAACAAATATACGCAACATCAATATACACCCCTGATCAACAAATAGAGAA 1980
|||
Db 1921 TACCAATCTCAACAAATATACGCAACATCAATATACACCCCTGATCAACAAATAGAGAA 1980
|||
Qy 1981 TCCTCAACGACACCA 1999
|||
Db 1981 TCCTCAACGACACCA 1999
|||

```

```

RESULT 2
US-08-961-083-227
; Sequence 227, Application US/08961083
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.

```

```

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 227:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 227:
US-08-961-083-227

```

```

Query Match 1.2%; Score 23; DB 4; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 TAAATCTACGACATATAAATC 23
|||
Db 11 TAAATCTACGACATATAAATC 33
|||

```

```

RESULT 3
US-09-540-209B-4827/c
; Sequence 4827, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 4827
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: B. fragilis
; US-09-540-209B-4827

```

```

Query Match 1.1%; Score 21; DB 5; Length 1389;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 986 TTGAGTACGGTGTCTACGAT 1006
|||
Db 729 TTGAGTACGGTGTCTACGAT 709
|||

```

```
RESULT 4
US-09-721-544-12893/C
; Sequence 12893, Application US/09721544
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jensen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fel
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/721,544
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 09/515,128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034,341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12893
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(332)
; OTHER INFORMATION: n = A,T,C or G
US-09-721-544-12893

Query Match 1.0%; Score 20; DB 5; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1867 TAGTCACTTCACAGCTCA 1886
Db 243 TAGTCACTTCACAGCTCA 224

RESULT 5
US-10-027-632-240647
; Sequence 240647, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
```

```
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 240647
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-240647
```

```
Query Match 1.0%; Score 20; DB 6; Length 502;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 699 AACAGGCTATACCTACTC 718
Db 147 aaacagctataactactc 166
```

```
RESULT 6
US-10-027-632-240648
; Sequence 240648, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 240648
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-240648
```

```
Query Match 1.0%; Score 20; DB 6; Length 502;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 699 AACAGGCTATACCTACTC 718
Db 147 aaacagctataactactc 166
```

```
RESULT 7
US-10-027-632-214358
; Sequence 214358, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027, 632
; SOFTWARE: FastSeq for Windows Version 4.0
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 214358
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-214358
```

```
Query Match 1.0%; Score 20; DB 6; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 699 AAACAGGCTATACCTACTC 718
DB 444 aaacaggtataactacc 463

RESULT 8
US-09-053-375B-301
; Sequence 375B-301
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CION-006
; CURRENT APPLICATION NUMBER: US/09/053,375B
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 301
; LENGTH: 4508
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-053-375B-301
```

```
Query Match 1.0%; Score 20; DB 5; Length 4508;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 319 AAAGCAACCAAGCAAGAAA 338
DB 560 aaagcaaccacgaaga 579

RESULT 9
US-09-620-393B-634/C
```

```
; Sequence 634, Application US/09620393B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1068P
; CURRENT APPLICATION NUMBER: US/09/620,393B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9948
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 634
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..427
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..427
; OTHER INFORMATION: Ceres Seq. ID 1377155
US-09-620-393B-634
```

```
Query Match 1.0%; Score 19; DB 5; Length 427;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1720 TGAAGGAAGCAATCCAGAA 1738
DB 78 TGAAGGAAGCAATCCAGAA 60
```

```
RESULT 10
US-09-620-393B-8776/C
; Sequence 8776, Application US/09620393B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1068P
; CURRENT APPLICATION NUMBER: US/09/620,393B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9948
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 8776
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..428
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..428
; OTHER INFORMATION: Ceres Seq. ID 1407382
US-09-620-393B-8776
```

```
Query Match 1.0%; Score 19; DB 5; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1720 TGAAGGAAGCAATCCAGAA 1738
DB 78 TGAAGGAAGCAATCCAGAA 60
```

```
RESULT 11
US-10-027-632-98352
; Sequence 98352, Application US/10027632
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98352
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-98352
```

```
Query Match 1.0%; Score 19; DB 6; Length 507;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 29 ATGCTGACTGGGTCTG 47
Db 319 atgcgcacttgggtctg 337
```

```
RESULT 12
US-10-027-632-250569/c
; Sequence 250569, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250569
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-250569
```

```
Query Match 1.0%; Score 19; DB 6; Length 1079;
```

```
Best Local Similarity 100.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 460 GGCTGGAATGCTCAGCA 478
Db 696 GGCTGGAATGCTCAGCA 678
```

```
RESULT 13
US-10-027-632-250570/c
; Sequence 250570, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250570
; LENGTH: 1079
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-250570
```

```
Query Match 1.0%; Score 19; DB 6; Length 1079;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 460 GGCTGGAATGCTCAGCA 478
Db 696 GGCTGGAATGCTCAGCA 678
```

```
RESULT 14
US-10-027-632-250571/c
; Sequence 250571, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```


;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 250571
;; LENGTH: 1079
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-250571

Query Match 1.0%; Score 19; DB 6; Length 1079;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 460 GGCTGGAATGCTCAGCA 478
|||||
Db 696 GGCTGGAATGCTCAGCA 678

RESULT 15
US-60-360-039-23856
; Sequence 23856, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkley, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360.039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23856
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-60-360-039-23856

Query Match 1.0%; Score 19; DB 7; Length 1299;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 99 CAATGTTCTATCGAGA 117
|||||
Db 857 caatglttctatcgaaga 875

RESULT 16
US-60-360-039-24027
; Sequence 24027, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkley, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360.039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24027
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-60-360-039-24027

Query Match 1.0%; Score 19; DB 7; Length 1299;

Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 99 CAATGTTCTATCGAGA 117
|||||
Db 857 caatglttctatcgaaga 875

RESULT 17
US-60-360-039-42540
; Sequence 42540, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkley, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360.039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 42540
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Anabaena PCC7120
US-60-360-039-42540

Query Match 1.0%; Score 19; DB 7; Length 1920;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 184 GCAAGCAATTCCTCCAA 202
|||||
Db 923 gcaagcaatctcccca 941

RESULT 18
US-09-539-331D-25958
; Sequence 25958, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539.331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 25958
; LENGTH: 166
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00485767
US-09-539-331D-25958

Query Match 0.9%; Score 18; DB 5; Length 166;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1284 AAAAGATGCTGCTT 1301
|||||
Db 41 aaagatgctgctgtctt 58

```

RESULT 19
US-09-539-800C-2484
; Sequence 2484, Application US/09539800C
; GENERAL INFORMATION:
; APPLICANT: Sellhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250,003
; PRIOR FILING DATE: February 10, 1999
; PRIOR APPLICATION NUMBER: 60/074,364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452,747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 2484
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: hu00209576
US-09-539-800C-2484

Query Match 0.9%; Score 18; DB 5; Length 229;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 AAAAGTGGCTGCTGCTT 1301
DB 117 aaagatgctgctgctt 134

```

```

; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/005,526
; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/903,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894

```

Db 38 ATGACAGCCTATTCTAA 21

RESULT 22

US-09-539-800C-15921

Sequence 15921, Application US/09539800C

GENERAL INFORMATION:

APPLICANT: Selhamer, Jeffrey J.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Stuart, Susan G.

APPLICANT: Stuve, Laura L.

APPLICANT: Naughton, Rebecca E.

TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE

FILE REFERENCE: PD-1023 CTP

CURRENT APPLICATION NUMBER: US/09/539,800C

PRIOR FILING DATE: 2000-03-30

PRIOR APPLICATION NUMBER: 08/521,383

PRIOR FILING DATE: August 16, 1995

PRIOR APPLICATION NUMBER: 08/271,217

PRIOR FILING DATE: June 27, 1994

PRIOR APPLICATION NUMBER: 08/334,881

PRIOR FILING DATE: November 4, 1994

PRIOR APPLICATION NUMBER: 08/943,978

PRIOR FILING DATE: October 3, 1997

PRIOR APPLICATION NUMBER: 60/028,732

PRIOR FILING DATE: October 4, 1996

PRIOR APPLICATION NUMBER: 08/943,979

PRIOR FILING DATE: October 4, 1997

PRIOR APPLICATION NUMBER: 60/027,782

PRIOR FILING DATE: October 4, 1996

PRIOR APPLICATION NUMBER: 08/993,774

PRIOR FILING DATE: December 18, 1997

PRIOR APPLICATION NUMBER: 60/034,975

PRIOR FILING DATE: December 20, 1996

PRIOR APPLICATION NUMBER: 09/250,003

PRIOR FILING DATE: February 10, 1999

PRIOR APPLICATION NUMBER: 60/074,364

PRIOR FILING DATE: February 12, 1998

PRIOR APPLICATION NUMBER: 09/452,747

PRIOR FILING DATE: December 1, 1999

PRIOR APPLICATION NUMBER: 60/111,910

PRIOR FILING DATE: December 10, 1998

NUMBER OF SEQ ID NOS: 19698

SOFTWARE: PERL Program

SEQ ID NO: 15921

LENGTH: 235

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No: hu01253724

US-09-539-800C-15921

Query Match 0.9%; Score 18; DB 5; Length 250;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 AAAAGATGCTGCTGCTT 1301

Db 117 aaaagatgctgctgctt 134

RESULT 21

US-09-539-331D-33884/C

Sequence 33884, Application US/09539331D

GENERAL INFORMATION:

APPLICANT: Selhamer, Jeffrey J.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Stuart, Susan G.

APPLICANT: Stuve, Laura L.

APPLICANT: Mullahy, Sara J.

APPLICANT: Naughton, Rebecca E.

TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE

FILE REFERENCE: PD-1022 CTP

CURRENT APPLICATION NUMBER: US/09/539,331D

PRIOR FILING DATE: 2000-03-30

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 40961

SOFTWARE: PERL Program

SEQ ID NO: 33884

LENGTH: 250

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No: hu01166483

US-09-539-331D-33884

Query Match 0.9%; Score 18; DB 5; Length 250;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1627 ATGACAGCCTATTCTAA 1644

|||||

Db 38 ATGACAGCCTATTCTAA 21

RESULT 23

US-09-539-800C-16898

Sequence 16898, Application US/09539800C

GENERAL INFORMATION:

APPLICANT: Selhamer, Jeffrey J.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Stuart, Susan G.

APPLICANT: Stuve, Laura L.

APPLICANT: Mullahy, Sara J.

```
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09/539,800C
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 08/521,383
; PRIOR FILING DATE: August 16, 1995
; PRIOR APPLICATION NUMBER: 08/271,217
; PRIOR FILING DATE: June 27, 1994
; PRIOR APPLICATION NUMBER: 08/334,881
; PRIOR FILING DATE: November 4, 1994
; PRIOR APPLICATION NUMBER: 08/943,978
; PRIOR FILING DATE: October 3, 1997
; PRIOR APPLICATION NUMBER: 60/028,732
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/943,979
; PRIOR FILING DATE: October 4, 1997
; PRIOR APPLICATION NUMBER: 60/027,782
; PRIOR FILING DATE: October 4, 1996
; PRIOR APPLICATION NUMBER: 08/993,774
; PRIOR FILING DATE: December 18, 1997
; PRIOR APPLICATION NUMBER: 60/034,975
; PRIOR FILING DATE: December 20, 1996
; PRIOR APPLICATION NUMBER: 09/250,003
; PRIOR FILING DATE: February 10, 1999
; PRIOR APPLICATION NUMBER: 60/074,364
; PRIOR FILING DATE: February 12, 1998
; PRIOR APPLICATION NUMBER: 09/452,747
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: 60/111,910
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 16898
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu01110969
US-09-539-800C-16898

Query Match          0.9%; Score 18; DB 5; Length 264;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 AAAAGATGCTGCTGCTT 1301
Db 73 aaaagatgctgctgctt 90

RESULT 24
US-09-540-210B-24058
; SEQUENCE 24058, Application US/09/540,210B
; GENERAL INFORMATION:
; APPLICANT: Sellhammer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Muliahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF URINARY TRACT TISSUE
; FILE REFERENCE: PD-1037 CIP
; CURRENT APPLICATION NUMBER: US/09/540,210B
; CURRENT FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: 08/972,899
; PRIOR FILING DATE: November 18, 1997
; PRIOR APPLICATION NUMBER: 08/395,244
; PRIOR FILING DATE: February 27, 1995
; PRIOR APPLICATION NUMBER: 08/722,922
; PRIOR FILING DATE: September 27, 1996
; PRIOR APPLICATION NUMBER: 60/005,526

; PRIOR FILING DATE: September 29, 1995
; PRIOR APPLICATION NUMBER: 08/824,029
; PRIOR FILING DATE: March 25, 1997
; PRIOR APPLICATION NUMBER: 60/014,010
; PRIOR FILING DATE: March 25, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/903,555
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/862,178
; PRIOR FILING DATE: May 22, 1997
; PRIOR APPLICATION NUMBER: 60/018,217
; PRIOR FILING DATE: May 23, 1996
; PRIOR APPLICATION NUMBER: 08/881,589
; PRIOR FILING DATE: June 24, 1997
; PRIOR APPLICATION NUMBER: 60/021,275
; PRIOR FILING DATE: June 25, 1996
; PRIOR APPLICATION NUMBER: 08/903,802
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/023,308
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/905,881
; PRIOR FILING DATE: August 1, 1997
; PRIOR APPLICATION NUMBER: 60/025,204
; PRIOR FILING DATE: August 1, 1996
; PRIOR APPLICATION NUMBER: 08/903,471
; PRIOR FILING DATE: July 30, 1997
; PRIOR APPLICATION NUMBER: 60/025,478
; PRIOR FILING DATE: July 31, 1996
; PRIOR APPLICATION NUMBER: 08/903,556
; PRIOR FILING DATE: July 31, 1997
; PRIOR APPLICATION NUMBER: 60/025,217
; PRIOR FILING DATE: August 22, 1996
; PRIOR APPLICATION NUMBER: 08/937,142
; PRIOR FILING DATE: September 23, 1997
; PRIOR APPLICATION NUMBER: 60/026,598
; PRIOR FILING DATE: September 24, 1996
; PRIOR APPLICATION NUMBER: 08/960,746
; PRIOR FILING DATE: October 29, 1997
; PRIOR APPLICATION NUMBER: 60/030,144
; PRIOR FILING DATE: October 30, 1996
; PRIOR APPLICATION NUMBER: 08/826,847
; PRIOR FILING DATE: April 10, 1997
; PRIOR APPLICATION NUMBER: 60/015,533
; PRIOR FILING DATE: April 10, 1996
; PRIOR APPLICATION NUMBER: 08/755,524
; PRIOR FILING DATE: November 22, 1996
; PRIOR APPLICATION NUMBER: 60/007,495
; PRIOR FILING DATE: November 22, 1995
; PRIOR APPLICATION NUMBER: 09/021,031
; PRIOR FILING DATE: February 10, 1998
; PRIOR APPLICATION NUMBER: 60/039,325
; PRIOR FILING DATE: February 13, 1997
; PRIOR APPLICATION NUMBER: 09/035,172
; PRIOR FILING DATE: March 4, 1998
; PRIOR APPLICATION NUMBER: 60/040,431
; PRIOR FILING DATE: March 5, 1997
; PRIOR APPLICATION NUMBER: 09/041,894
; PRIOR FILING DATE: March 12, 1998
; PRIOR APPLICATION NUMBER: 60/040,199
; PRIOR FILING DATE: March 14, 1997
; PRIOR APPLICATION NUMBER: 09/050,817
; PRIOR FILING DATE: March 30, 1998
; PRIOR APPLICATION NUMBER: 60/043,792
; PRIOR FILING DATE: April 11, 1997
; PRIOR APPLICATION NUMBER: 09/074,999
; PRIOR FILING DATE: May 8, 1998
; PRIOR APPLICATION NUMBER: 60/048,431
; PRIOR FILING DATE: May 29, 1997
```

```
; PRIOR APPLICATION NUMBER: 09/107,592
; PRIOR FILING DATE: June 30, 1998
; PRIOR APPLICATION NUMBER: 60/052,751
; PRIOR FILING DATE: July 1, 1997
; PRIOR APPLICATION NUMBER: 09/094,079
; PRIOR FILING DATE: June 9, 1998
; PRIOR APPLICATION NUMBER: 60/049,975
; PRIOR FILING DATE: June 13, 1997
; NUMBER OF SEQ ID NOS: 35654
; SOFTWARE: PERL Program
; SEQ ID NO 24058
; LENGTH: 268
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00169618
US-09-540-210B-24058
```

```
Query Match          0.9%; Score 18; DB 5; Length 268;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1585 TGAAGTATTGCTGCTA 1602
Db 74 tgacctattgtcgcgcta 91
```

```
RESULT 25
US-09-919-002-7641
; Sequence 7641, Application US/09919002
; GENERAL INFORMATION:
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Liu, Jin
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/919,002
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922
; PRIOR FILING DATE: FILING DATE: 1999-07-22
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341
; PRIOR FILING DATE: FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7641
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-002-7641
```

```
Query Match          0.9%; Score 18; DB 5; Length 277;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1783 ATTTAAATGCTGCTCG 1800
Db 4 atttaaaatgctgcctcg 21
```

```
RESULT 26
US-09-539-331D-21402
; Sequence 21402, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
```

```
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 21402
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu00318340
; NAME/KEY: unsure
; LOCATION: 276, 281
; OTHER INFORMATION: a, t, c, g, or other
US-09-539-331D-21402
```

```
Query Match          0.9%; Score 18; DB 5; Length 285;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1284 AAAAGATGCTGCTGCTT 1301
Db 53 aaaagatgctgctgctt 70
```

```
RESULT 27
US-09-539-331D-34826
; Sequence 34826, Application US/09539331D
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CARDIOVASCULAR SYSTEM TISSUE
; FILE REFERENCE: PD-1022 CIP
; CURRENT APPLICATION NUMBER: US/09/539,331D
; CURRENT FILING DATE: 2000-03-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 40961
; SOFTWARE: PERL Program
; SEQ ID NO 34826
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: hu01325546
US-09-539-331D-34826
```

```
Query Match          0.9%; Score 18; DB 5; Length 296;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1284 AAAAGATGCTGCTGCTT 1301
Db 197 aaaagatgctgctgctt 214
```

```
RESULT 28
US-10-027-632-94603
; Sequence 94603, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
```

```
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94603
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-94603
```

```
Query Match      0.9%; Score 18; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 TTCCTCCCAAGTGTGATC 210
      |||||||
DB 436 ttccctcccaagtgatc 453
```

```
RESULT 29
US-10-027-632-94604
; Sequence 94604, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 94604
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-94604
```

```
Query Match      0.9%; Score 18; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 TTCCTCCCAAGTGTGATC 210
      |||||||
```

```
DB 436 ttccctcccaagtgatc 453
```

```
RESULT 30
US-10-027-632-305498
; Sequence 305498, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305498
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305498
```

```
Query Match      0.9%; Score 18; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 193 TTCCTCCCAAGTGTGATC 210
      |||||||
DB 436 ttccctcccaagtgatc 453
```

```
RESULT 31
US-10-027-632-305499
; Sequence 305499, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305499
; LENGTH: 457
```

```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305499

Query Match      0.9%; Score 18; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 TTCCTCCAGGTGATC 210
    |||
Db 436 ttcctccaagtgatc 453

RESULT 32
US-10-027-632-305500
; Sequence 305500, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 305500
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305500

Query Match      0.9%; Score 18; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 TTCCTCCAGGTGATC 210
    |||
Db 436 ttcctccaagtgatc 453

RESULT 33
US-10-027-632-286371/c
; Sequence 286371, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
```

```
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286371
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-286371

Query Match      0.9%; Score 18; DB 6; Length 478;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1530 ATACAGCAGGAATTG 1547
    |||
Db 201 ATACAGCAGGAATTG 184

RESULT 34
US-10-097-105-1069
; Sequence 1069, Application US/10097105
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: King, Gordon E.
; APPLICANT: Secrist, Heather
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.504C1
; CURRENT APPLICATION NUMBER: US/10/097,105
; CURRENT FILING DATE: 2002-03-13
; NUMBER OF SEQ ID NOS: 1562
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1069
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 318, 328, 347, 398, 400, 411, 423, 425, 429, 439, 443, 445,
; LOCATION: 458, 460, 478, 481, 498, 500, 502
; OTHER INFORMATION: n = A,T,C or G
US-10-097-105-1069

Query Match      0.9%; Score 18; DB 6; Length 512;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 AAAGATGCTGCTGCTT 1301
    |||
Db 181 aaagatgctgctgctt 198

RESULT 35
US-09-620-393B-6548
; Sequence 6548, Application US/09620393B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1068P
; CURRENT APPLICATION NUMBER: US/09/620,393B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9948
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6548
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..526
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 1..526
; OTHER INFORMATION: Ceres Seq. ID 1392936
US-09-620-393B-6548

Query Match
Best Local Similarity 100.0%; Score 18; DB 5; Length 526;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1268 TATGAGCAAGTAGTGA 1285
|||||
DB 248 tatgagcaagtagtgaa 265

RESULT 36
US-10-027-632-135305
; Sequence 135305, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 135305
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-135305

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 532;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1822 TCCACACACCCCATC 1839
|||||
DB 74 tccacacaccccatc 91

RESULT 37
US-10-027-632-81926/c
; Sequence 81926, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81926
; LENGTH: 581
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-81926

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 581;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 687 AAGTTGAGAGAAACAG 704
|||||
DB 95 AAGTTGAGAGAAACAG 78

RESULT 38
US-10-027-632-157285/c
; Sequence 157285, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157285
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-157285

Query Match
Best Local Similarity 100.0%; Score 18; DB 6; Length 850;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 966 TCACAGACTATGCTCTG 983
|||||
Db 635 TCACAGACTATGCTCTG 618

RESULT 39
US-10-106-698-791
; Sequence 791, Application US/10106698
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentln Ver. 3.0
; SEQ ID NO 791
; LENGTH: 1229
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (19)..(19)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (305)..(305)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-791

Query Match 0.9%; Score 18; DB 6; Length 1229;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1627 ATGACAGGCTATCTAA 1644
|||||
Db 779 atgacagcgtatctctaa 796

RESULT 40
US-09-935-625-18639
; Sequence 18639, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 18639
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: Misc.feature
; LOCATION: 1..1281
; OTHER INFORMATION: Ceres Seq. ID no. 1015072
US-09-935-625-18639

Query Match 0.9%; Score 18; DB 5; Length 1281;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1971 AAATCAGATCCTCAAC 1988
|||||

Db 676 aaatcagaatcctcaac 693

RESULT 41
US-09-935-625-27058
; Sequence 27058, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 27058
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: Misc.feature
; LOCATION: 1..1281
; OTHER INFORMATION: Ceres Seq. ID no. 1015072
US-09-935-625-27058

Query Match 0.9%; Score 18; DB 5; Length 1281;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1971 AAATCAGATCCTCAAC 1988
|||||
Db 676 aaatcagaatcctcaac 693

RESULT 42
US-60-360-039-47071
; Sequence 47071, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 47071
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-60-360-039-47071

Query Match 0.9%; Score 18; DB 7; Length 1293;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 646 CCCTGCTTACATGATTA 663
|||||
Db 399 ccctgcttacatgataa 416

RESULT 43
US-60-360-039-47214
; Sequence 47214, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)A
CURRENT APPLICATION NUMBER: US/60/360,039
CURRENT FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 47214
LENGTH: 1299
TYPE: DNA
ORGANISM: Escherichia coli
US-60-360-039-47214

Query Match 0.9%; Score 18; DB 7; Length 1299;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 ATCGTTTCATCGAAGAC 118
|||||

DB 856 atcgctctatcgagac 873

RESULT 44
US-09-919-002-13066/c
Sequence 13066, Application US/09919002
GENERAL INFORMATION:
APPLICANT: Leshkowitz, Dena
APPLICANT: Liu, Jin
TITLE OF INVENTION: NOVEL CONFIGS OBTAINED FROM VARIOUS CDNA
TITLE OF INVENTION: LIBRARIES
FILE REFERENCE: 20411-752CON1
CURRENT APPLICATION NUMBER: US/09/919,002
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922
PRIOR FILING DATE: FILING DATE: 1999-07-22
PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341
PRIOR FILING DATE: FILING DATE: 1998-02-13
NUMBER OF SEQ ID NOS: 13203
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13066
LENGTH: 1328
TYPE: DNA
ORGANISM: Homo sapiens
US-09-919-002-13066

Query Match 0.9%; Score 18; DB 5; Length 1328;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 AAAAGATGCTGCTT 1301
|||||

DB 525 AAAAGATGCTGCTT 508

RESULT 45
US-10-027-632-253001/c
Sequence 253001, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 253001
LENGTH: 1555
TYPE: DNA
ORGANISM: Human
US-10-027-632-253001

Query Match 0.9%; Score 18; DB 6; Length 1555;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 AATGGAACATATGGAATG 391
|||||

DB 704 AATGGAACATATGGAATG 687

RESULT 46
US-10-027-632-253002/c
Sequence 253002, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 253002
LENGTH: 1555
TYPE: DNA
ORGANISM: Human
US-10-027-632-253002

Query Match 0.9%; Score 18; DB 6; Length 1555;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 AATGGAACATATGGAATG 391
|||||

DB 704 AATGGAACATATGGAATG 687

RESULT 47
US-10-027-632-253003/c
Sequence 253003, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

;; TITLE OF INVENTION: Polymorphisms in the Human Genome
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 253003
;; LENGTH: 1555
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-253003

Query Match 0.9%: Score 18; DB 6; Length 1555;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 374 AATGGAGCTATGATG 391
|||||
Db 704 AATGGAGCTATGATG 687

RESULT 48
US-10-027-632-253004/C
;; Sequence 253004, Application US/10027632
;; GENERAL INFORMATION:
;; APPLICANT: Wang, David G.
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;; FILE REFERENCE: 108827.129
;; CURRENT APPLICATION NUMBER: US/10/027,632
;; CURRENT FILING DATE: 2002-04-30
;; PRIOR APPLICATION NUMBER: US 60/218,006
;; PRIOR FILING DATE: 2000-07-12
;; PRIOR APPLICATION NUMBER: US 60/198,676
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: US 60/193,483
;; PRIOR FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: US 60/185,218
;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 253004
;; LENGTH: 1555
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-253004

Query Match 0.9%: Score 18; DB 6; Length 1555;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 374 AATGGAGCTATGATG 391
|||||
Db 704 AATGGAGCTATGATG 687

RESULT 49
US-60-360-039-24796/C
;; Sequence 24796, Application US/60360039
;; GENERAL INFORMATION:
;; APPLICANT: Cao, Yongwei
;; APPLICANT: Chen, Xianfeng
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)A
;; CURRENT APPLICATION NUMBER: US/60/360,039
;; CURRENT FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO: 24796
;; LENGTH: 1773
;; TYPE: DNA
;; ORGANISM: Methanobacterium thermoautotrophicum
US-60-360-039-24796

Query Match 0.9%: Score 18; DB 7; Length 1773;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1888 CTCACCACTCCACGAC 1905
|||||
Db 837 CTCACCACTCCACGAC 820

RESULT 50
US-10-115-123-53
;; Sequence 53, Application US/10115123
;; GENERAL INFORMATION:
;; APPLICANT: Ruben et al.
;; TITLE OF INVENTION: 94 Human Secreted Proteins
;; FILE REFERENCE: P2029630AIPD2
;; CURRENT APPLICATION NUMBER: US/10/115,123
;; CURRENT FILING DATE: 2002-04-04
;; PRIOR APPLICATION NUMBER: PCT/US99/13418
;; PRIOR FILING DATE: 1999-06-15
;; PRIOR APPLICATION NUMBER: 60/089,507
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089,508
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089,509
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089,510
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/090,112
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090,113
;; PRIOR FILING DATE: 1998-06-22
;; NUMBER OF SEQ ID NOS: 532
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 53
;; LENGTH: 2288
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (940)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: SITE
;; LOCATION: (1279)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: SITE

```
;
; LOCATION: (1798)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2280)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2285)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-115-123-53
```

```
Query Match          0.9%; Score 18; DB 6; Length 2288;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1284 AAAAGATGCTGCTGCTT 1301
    ||||||||||||||||
Db 1734 aaaagatgctgctgctt 1751
```

Search completed: June 13, 2002, 13:44:17
Job time: 7794 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:40:34 ; Search time 21.48 Seconds
(without alignments)
2979.306 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KIYDNKNQIADLGSESRVYN.....TQSSNTTPOQNONQPPAQP 666

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	719	2	G95042 penicillin-binding
2	3475	99.7	719	2	S28036 penicillin-binding
3	3471	99.6	719	2	S28037 penicillin-binding
4	3466	99.5	719	2	A42893 penicillin-binding
5	3466	99.5	719	2	A97913 peptidoglycan glyc
6	3246	93.2	719	2	S28034 penicillin-binding
7	3245	93.1	719	2	S28032 penicillin-binding
8	3243	93.1	719	2	S28031 penicillin-binding
9	3241	93.0	719	2	S28033 penicillin-binding
10	3192	91.6	719	2	S28035 penicillin-binding
11	2820	80.9	608	2	S28036 penicillin-binding
12	2698.5	77.5	637	2	B42893 penicillin-binding
13	1657.5	47.6	664	2	G86692 penicillin-binding
14	1217	34.9	827	2	AD1311 penicillin-binding
15	1201	34.5	826	2	AD1683 penicillin-binding
16	1074.5	30.8	886	2	F83862 penicillin-binding
17	999	28.7	914	2	I40529 penicillin-binding
18	914	26.2	188	2	S31952 penicillin-binding
19	817.5	23.5	727	2	B89923 penicillin-binding
20	813.5	23.3	716	2	S43693 penicillin-binding
21	802.5	23.0	809	2	F97183 penicillin-binding
22	739.5	21.2	714	2	AE1353 membrane carboxype
23	739	21.2	714	2	AE1353 penicillin-binding
24	724	20.8	714	2	AE1723 penicillin-binding
25	712	20.4	726	2	F70355 penicillin-binding
26	698.5	20.0	719	2	A83800 penicillin-binding
27	670.5	19.2	643	2	EB1396 penicillin-binding
28	656	18.8	966	2	EB4053 penicillin-binding
29	655	18.8	764	2	B97371 penicillin-binding

30	655	18.8	764	2	AB2589 penicillin-binding
31	649.5	18.6	743	2	D86888 penicillin-binding
32	642.5	18.4	822	2	F83016 penicillin-binding
33	640	18.4	798	2	H81040 penicillin-binding
34	637	18.3	731	2	D95235 penicillin-binding
35	636	18.3	731	2	E98099 peptidoglycan glyc
36	627	18.0	659	2	E64594 penicillin-binding
37	625	17.9	660	2	F71917 penicillin-binding
38	618	17.7	851	2	AF0018 penicillin-binding
39	617	17.7	835	2	B82051 penicillin-binding
40	615	17.7	735	2	AI2174 penicillin-binding
41	614	17.6	873	2	B75514 penicillin-binding
42	613.5	17.6	687	2	D84126 penicillin-binding
43	608	17.5	718	2	AI3420 penicillin-binding
44	598.5	17.2	624	2	A55220 penicillin-binding
45	598	17.2	885	2	S76357 penicillin-binding

ALIGNMENTS

RESULT 1

G95042 penicillin-binding protein 1A [imported] - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001

C:Accession: G95042

R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf

erson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris

A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A:Reference number: A95000, MUID:21357209, PMID:11463916

A:Accession: G95042

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-719 <CUR>

A:Cross-references: GB:AE005672; PIDN:AAK74536.1; PID:gl4971838; GSPDB:GN00164; TIGR:

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP0369

C:Superfamily: penicillin-binding protein 1B

Query Match Best Local Similarity 100.0%; Score 3484; DB 2; Length 719;

Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KIYDNKNQIADLGSESRVYNQANDIPTDLVKAIVSIEDHRRFDRGIDTIRIIGAFARN	60
DB	54	KIYDNKNQIADLGSESRVYNQANDIPTDLVKAIVSIEDHRRFDRGIDTIRIIGAFARN	113
QY	61	LOSNSLOGSTLTQOLIKLTFYFSTSDQTSRKQAEAMLAIOEQKATKOEILTYVINK	120
DB	114	LOSNSLOGSTLTQOLIKLTFYFSTSDQTSRKQAEAMLAIOEQKATKOEILTYVINK	173
QY	121	VYNSNGNGMGTAAQNTYGRKLNLSIFQALLAGMPAPAPQYDPSHPEAAQDRNLVL	180
DB	174	VYNSNGNGMGTAAQNTYGRKLNLSIFQALLAGMPAPAPQYDPSHPEAAQDRNLVL	233
QY	181	SEMNQGYTSAEYERKAVNPITDGLQSLKASVNPAPMDVYLFKEVINQVEEETGYNLLT	240
DB	234	SEMNQGYTSAEYERKAVNPITDGLQSLKASVNPAPMDVYLFKEVINQVEEETGYNLLT	293
QY	241	TGMDVYTNVDEQAKHMLDIYNTDEYVAYPDDELQVASTIVDVSNGKYIAQLGARHSSN	300
DB	294	TGMDVYTNVDEQAKHMLDIYNTDEYVAYPDDELQVASTIVDVSNGKYIAQLGARHSSN	353
QY	301	VSFGINQAVETNRDQSGTMRKITDYAPALEGYVDSTATYVHDEPYNPGNTPTVYNNDR	360
DB	354	VSFGINQAVETNRDQSGTMRKITDYAPALEGYVDSTATYVHDEPYNPGNTPTVYNNDR	413

OY	361	GYESNITTOLOALOOSSRNPAVELLNKVGILNRKPTFLNGIDIPSIHNSMISSTTSSD	420
Dd	414	GIFGNITTOLOALOOSSRNPAVELLNKVGILNRKPTFLNGIDIPSIHNSMISSTTSSD	473
OY	421	KRYGASSEKMAAAYAFANGTTYKPMYIHKVFSDSGEKEFSNVGTBRAMETTAYMTD	480
Dd	474	KRYGASSEKMAAAYAFANGTTYKPMYIHKVFSDSGEKEFSNVGTBRAMETTAYMTD	533
OY	481	MKKTVLTJTGTSNAALFAMLPOAGKTGTSNYTDEETLENIHKTSOFAPBELRAGYRRKSM	540
Dd	534	MKKTVLTJTGTSNAALFAMLPOAGKTGTSNYTDEETLENIHKTSOFAPBELRAGYRRKSM	593
OY	541	AWMTGYSNRLPLVLVNGLTLAAKYTRSMYTYISEGSNPEDMNIEPGLYRNGEBFYFKCAR	600
Dd	594	AWMTGYSNRLPLVLVNGLTLAAKYTRSMYTYISEGSNPEDMNIEPGLYRNGEBFYFKCAR	653
OY	601	STMWNPAPQOPSTESSSSSDSSTSOSSSTTPSTNNSTTNPNNNNTQQOQNTTPOOQN	660
Dd	654	STMWNPAPQOPSTESSSSSDSSTSOSSSTTPSTNNSTTNPNNNNTQQOQNTTPOOQN	713
OY	661	PQPAPQ 666 	
Dd	714	PQPAPQ 719	

RESULT 2
S28038
penicillin-binding protein Ia - Streptococcus pneumoniae (strain 45607) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28038
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein Ia genes from different clones of penicillin-resistant strains of *Streptococcus pneumoniae*
A:Reference number: S28031; MUID:93010977
A:Accession: S28038
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67873; NID:q47419; PIDN:CAA#8073.1; PID:q47420
C:Superfamily: penicillin-binding protein IB

Query Match	99.7%	Score 3475	DB 2	Length 719
Best Local Similarity	99.7%	Pred. No. 1.5e-193		
Matches 664	Conservative 2	Missmatches 0	Indels 0	Gaps 0
OY	1 KTYDNKNQIADLGSERRNAQAANIPTLVYVAIYSIEDHREFDRHGIDITRILGAFUEN	60		
Db	54 KTYDNKNQIADLGSERRNAQAANIPTLVYVAIYSIEDHREFDRHGIDITRILGAFUEN	113		
OY	61 LQSNLSGGSTLTQOLIKLTFPSTSDQTSIRKQAEAMLAQLQKQKATQOEILTYIYNK	120		
Db	114 LQSNLSGGSTLTQOLIKLTFPSTSDQTSIRKQAEAMLAQLQKQKATQOEILTYIYNK	173		
OY	121 VYMSNGTGMQTAQAONYGKDLNNLSLPOLALLAGPQAPNOYDPYSHPDEAADRRLTVL	180		
Db	174 VYMSNGTGMQTAQAONYGKDLNNLSLPOLALLAGPQAPNOYDPYSHPDEAADRRLTVL	233		
OY	181 SEMKNOGYSIAQYRKAVNPTITDQLQSLKASNPAYMDNLKVEYINQVEEETGYNLT	240		
Db	234 SEMKNOGYSIAQYRKAVNPTITDQLQSLKASNPAYMDNLKVEYINQVEEETGYNLT	293		
OY	241 TGMADVTVNDDQAKHLMDIYNTDEYVAPDDELQVASTIYVNSGKYTAQLAGARHOSSN	300		
Db	294 TGMADVTVNDDQAKHLMDIYNTDEYVAPDDELQVASTIYVNSGKYTAQLAGARHOSSN	353		
OY	301 VSFQINGQAVETTRDQGSTMKPPTDAPAPALEYGVYSTATIVHDEPYNPQGTNPYNNMR	360		
Db	354 VSFQINGQAVETTRDQGSTMKPPTDAPAPALEYGVYSTATIVHDEPYNPQGTNPYNNMR	413		
OY	361 GYFGNITLQYALQOSRNPVAVETLNKVGINRAKTFPLNGIGIDYPSIHSNAISSNTTSED	420		

Db	414	GYGCHITTLQYALQOOSRNVPAVELINKGILMRATFLNGLOIDVPSLHYNSAITSNTEED	473
QY	421	KKYGASSKKMAAAYAAFPANGTTYKKPYIHKVYPSDGEKESNVGTGRAMKETTAYMATD	480
Db	474	KKYGASSSEKKMAAAAPANGTTYKKPYIHKVYFSDSEKEFNVGTGRAMKETTAYMATME	533
QY	481	MMKTYLTLYGTGRNMYLAMLPOACKTGTSNTDEIEIENHITSOFVAPDELFGYTRKYSM	540
Db	534	MMKTYLTLYGTGRNMYLAMLPOACKTGTSNTDEIEIENHITSOFVAPDELFGYTRKYSM	593
QY	541	AWMTGYSNRKLTPVLVNGNLITAAAKYIKRSMNTYYLSEGSNPEDMNITPEGLYRNGEYVFENKAR	600
Db	594	AWMTGYSNRKLTPVLVNGNLITAAAKYIKRSMNTYYLSEGSNPEDMNITPEGLYRNGEYVFENKAR	653
QY	601	STWNSPAQOQPPSTESSSSSDSSSTQSSSTPESTNNSTTTNNNNNTQOOSNTTTPDOONON	660
Db	654	STWNSPAQOQPPSTESSSSSDSSSTQSSSTPESTNNSTTTNNNNNTQOOSNTTTPDOONON	713
QY	661	POPAPQ 666	
Db	714	POPAPQ 719	

RESULT 3
S28037
penicillin-binding protein Ia - Streptococcus pneumoniae (strain 63915) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28037
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein Ia genes from different clones of
A:Reference number: S28031; MUID:93010937
A:Accession: S28037
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67872; NID:g47417; PIDN:CAA48072.1; PID:g47418
C:Superfamily: penicillin-binding protein IB

Query Match	99.6%	Score 3471	DB 2	Length 719
Best Local Similarity	99.4%	Pred. No. 2,6e193		
Matches 662	Conservative	4	Mismatches 0	Indels 0
				Gaps 0
QY	1	KIYDKNKQIADLGSERRVNAQANDIPMDLVVAIYSIEDHREFDRHGRDITIRILGAFLEKN	60	
DB	54	KIYDKNKQIADLGSERRVNAQANDIPMDLVVAIYSIEDHREFDRHGRDITIRILGAFLEKN	113	
QY	61	LOSNSLOGSGSTLTQOLIKLITYFSTSTSDQTSIRKQOEAMLAIQLEOKATKQIILTYINK	120	
DB	114	LOSNSLOGSGSTLTQOLIKLITYFSTSTSDQTSIRKQOEAMLAIQLEOKATKQIILTYINK	173	
QY	121	YVMSNGNNGMOTPAANOYNGKDLNNLSLPOLALLAGMPQAPNOYDEYSHPHEAODRRNLVL	180	
DB	174	YVMSNGNNGMOTPAANOYNGKDLNNLSLPOLALLAGMPQAPNOYDEYSHPHEAODRRNLVL	233	
QY	181	SEMNKOGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNLYKEVINOVEEETGYNLT	240	
DB	234	SEMNKOGYISAEQYERKAVNPTITDGLQSLKSASNPAYMDNLYKEVINOVEEETGYNLT	293	
QY	241	TGMDVYTVNDQEAOKHMDIYNTDEYVAVPDELOVASTIVVNSNGKYIAOLGARRHOSN	300	
DB	294	TGMDVYTVNDQEAOKHMDIYNTDEYVAVPDELOVASTIVVNSNGKYIAOLGARRHOSN	353	
QY	301	VSFGINQAVETNRDMSGSTMKPTITDPAALAEVGVYDSTATIVHDEBYNPGTTPVYNNDR	360	
DB	354	VSFGINQAVETNRDMSGSTMKPTITDPAALAEVGVYDSTATIVHDEBYNPGTTPVYNNDR	413	
QY	361	GYFGNITLQYALQOOSRRNPAYVETLKKVGLGNRAKKTFLNGLGIDYPSIHTSNALSSNTTESD	420	
DB	414	GYFGNITLQYALQOOSRRNPAYVETLKKVGLGNRAKKTFLNGLGIDYPSIHTSNALSSNTTESD	473	

QY 421 KKGASSEKMAAAYAAAFANGGTYKKPMYIHKVPSDGSSEKESNVGTRAMKETTAYMTD 480
|||||
Db 474 KKGASSEKMAAAYAAAFANGGTYKKPMYIHKVPSDGSSEKESNVGTRAMKETTAYMTD 533
QY 481 MKKTYLTGTGRNAYLAAPGAGTGTSTNTDEELENHIKTSQYAPDELFAGTTRKXSM 540
|||||
Db 534 MKKTYLTGTGRNAYLAAPGAGTGTSTNTDEELENHIKTSQYAPDELFAGTTRKXSM 593
QY 541 AVMTGYSNRLTPLYGNGLTVAAYKYSRMMTYLSEGSNPEDMNIPGGLYRNGEFPKNGAR 600
|||||
Db 594 AVMTGYSNRLTPLYGNGLTVAAYKYSRMMTYLSEGSNPEDMNIPGGLYRNGEFPKNGAR 653
QY 601 STWNSPAPQPPSTESSSSSDSSSTQSSSTPTSTNNSTTNPNNNTQOQSTTTPDOQON 660
|||||
Db 654 STWNSPAPQPPSTESSSSSDSSSTQSSSTPTSTNNSTTNPNNNTQOQSTTTPDOQON 713
QY 661 POPAP 666
|||||
Db 714 POPAP 719

RESULT 4

A42893
penicillin-binding protein 1A - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Jan-2000
C:Accession: A42893
R:Martin, C.; Briese, T.; Hakenbeck, R.
J. Bacteriol. 174, 4517-4523, 1992
A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae.
A:Reference number: A42893; MUID:92325042
A:Accession: A42893
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: GB:M90527; NID:g153766; PIDN:AAA26956.1; PID:g153768
C:Superfamily: penicillin-binding protein 1B

Query Match 99.5%; Score 3466; DB 2; Length 719;
Best Local Similarity 99.4%; Pred. No. 5.1e-193;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSERRVNAQANDIPTDLKAYSIDHRRFPDHRGIDTRIIGAFLRN 60
|||||
Db 54 KIYDNKNLIDLGSERRVNAQANDIPTDLKAYSIDHRRFPDHRGIDTRIIGAFLRN 113
QY 61 LQSNLSGGSTLTQOLIKLTYFSTSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 120
|||||
Db 114 LQSNLSGGSTLTQOLIKLTYFSTSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 173
QY 121 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPQAPQYDPSHPEAAQDRNLVL 180
|||||
Db 174 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPQAPQYDPSHPEAAQDRNLVL 233
QY 181 SEMKNOCYTSAEYKAVNPITDGLQSLKASNPAYAMDYKLEVINQVEEFGYNLLT 240
|||||
Db 234 SEMKNOCYTSAEYKAVNPITDGLQSLKASNPAYAMDYKLEVINQVEEFGYNLLT 293
QY 241 TGMVYTNVDEQAKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 300
|||||
Db 294 TGMVYTNVDEQAKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 353
QY 301 VSEGINQAVETNRDWSGTMKPTIDYAPALEGYVDSTATTYHDEPYNPGTNTPYNMNR 360
|||||
Db 354 VSEGINQAVETNRDWSGTMKPTIDYAPALEGYVDSTATTYHDEPYNPGTNTPYNMNR 413
QY 361 GYFGNITLQYVALQOOSRNPAVETLNKVGILNRAKTFNLGLIDYPSIHYSNAISSNTESD 420
|||||
Db 414 GYFGNITLQYVALQOOSRNPAVETLNKVGILNRAKTFNLGLIDYPSIHYSNAISSNTESD 473
QY 421 KKGASSEKMAAAYAAAFANGGTYKKPMYIHKVPSDGSSEKESNVGTRAMKETTAYMTD 480

Db 474 KKGASSEKMAAAYAAAFANGGTYKKPMYIHKVPSDGSSEKESNVGTRAMKETTAYMTD 533
|||||
QY 481 MKKTYLTGTGRNAYLAAPGAGTGTSTNTDEELENHIKTSQYAPDELFAGTTRKXSM 540
|||||
Db 534 MKKTYLTGTGRNAYLAAPGAGTGTSTNTDEELENHIKTSQYAPDELFAGTTRKXSM 593
QY 541 AVMTGYSNRLTPLYGNGLTVAAYKYSRMMTYLSEGSNPEDMNIPGGLYRNGEFPKNGAR 600
|||||
Db 594 AVMTGYSNRLTPLYGNGLTVAAYKYSRMMTYLSEGSNPEDMNIPGGLYRNGEFPKNGAR 653
QY 601 STWNSPAPQPPSTESSSSSDSSSTQSSSTPTSTNNSTTNPNNNTQOQSTTTPDOQON 660
|||||
Db 654 STWNSPAPQPPSTESSSSSDSSSTQSSSTPTSTNNSTTNPNNNTQOQSTTTPDOQON 713
QY 661 POPAP 666
|||||
Db 714 POPAP 719

RESULT 5

A97913
peptidoglycan glycosyltransferase (EC 2.4.1.129) [imported] - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: A97913
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnthen, S.; y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97913
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAK99133.1; PID:g15457886; GSPDB:GN00174
C:Genetics:
A:Gene: pbppA
C:Superfamily: penicillin-binding protein 1B
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 99.5%; Score 3466; DB 2; Length 719;
Best Local Similarity 99.4%; Pred. No. 5.1e-193;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSERRVNAQANDIPTDLKAYSIDHRRFPDHRGIDTRIIGAFLRN 60
|||||
Db 54 KIYDNKNLIDLGSERRVNAQANDIPTDLKAYSIDHRRFPDHRGIDTRIIGAFLRN 113
QY 61 LQSNLSGGSTLTQOLIKLTYFSTSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 120
|||||
Db 114 LQSNLSGGSTLTQOLIKLTYFSTSTSDQTSRKAQEWALAIQLEOKATKOEILTYINK 173
QY 121 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPQAPQYDPSHPEAAQDRNLVL 180
|||||
Db 174 VYMSNGNGMOTAAQNYGKDLNLSLPQALLAGMPQAPQYDPSHPEAAQDRNLVL 233
QY 181 SEMKNOCYTSAEYKAVNPITDGLQSLKASNPAYAMDYKLEVINQVEEFGYNLLT 240
|||||
Db 234 SEMKNOCYTSAEYKAVNPITDGLQSLKASNPAYAMDYKLEVINQVEEFGYNLLT 293
QY 241 TGMVYTNVDEQAKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 300
|||||
Db 294 TGMVYTNVDEQAKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKVIQALGARHOSN 353
QY 301 VSEGINQAVETNRDWSGTMKPTIDYAPALEGYVDSTATTYHDEPYNPGTNTPYNMNR 360
|||||
Db 354 VSEGINQAVETNRDWSGTMKPTIDYAPALEGYVDSTATTYHDEPYNPGTNTPYNMNR 413
QY 361 GYFGNITLQYVALQOOSRNPAVETLNKVGILNRAKTFNLGLIDYPSIHYSNAISSNTESD 420

```
|||||
Db 414 GYFGNTTLOYALQOOSNNVAVETLNKVGILNRAKTFPLNGLGIDIPSIHYSAISSNTTESD 473
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTTAYMTMD 480
Db 474 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTTAYMTMD 533
OY 481 MKKTVLTGTGRNAYLAMPLOAGKGTGTSNYTDEIEINHIKTSOFVAPDELFACTYRKYSK 540
Db 534 MKKTVLTGTGRNAYLAMPLOAGKGTGTSNYTDEIEINHIKTSOFVAPDELFACTYRKYSK 593
OY 541 AVWTGYSNRLTPLYGNGLTVAAKVYRSMTYIYSEGSNPEDMNIPEGLYRNGEFVFNKGAR 600
Db 594 AVWTGYSNRLTPLYGNGLTVAAKVYRSMTYIYSEGSNPEDMNIPEGLYRNGEFVFNKGAR 653
OY 601 STWNSPAPQOPPESTESSSSSDSTSSSTPTSTNNSTTTNNNTQOOSNTTPOOONON 660
Db 654 STWNSPAPQOPPESTESSSSSDSTSSSTPTSTNNSTTTNNNTQOOSNTTPOOONON 713
OY 661 POPAP 666
Db 714 POPAP 719
```

```
RESULT 6
S28034
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 56742) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28034
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of pen
A:Reference number: S28031; MUID:93010977
A:Accession: S28034
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-References: EMBL:X67869; NID:947411; PIDN:CAA48069.1; PID:947412
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 93.2%; Score 3246; DB 2; Length 719;
Best Local Similarity 92.0%; Pred. No. 2,8e-180;
Matches 613; Conservative 31; Mismatches 22; Indels 0; Gaps 0;
OY 1 KIYDNKNOLIADGSERRVNAQANDIPTDLVKAIVSIEDHRRFDHGIDIRILGAFLRN 60
Db 54 KIYDNKNOLIADGSERRVNAQANDIPTDLVKAIVSIEDHRRFDHGIDIRILGAFLRN 113
OY 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWMLAIQLEQKATKQETLTYIYINK 120
Db 114 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWMLAIQLEQKATKQETLTYIYINK 173
OY 121 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRNLVL 180
Db 174 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRNLVL 233
OY 181 SEMKNGYISADQYKAVNTPITDGLQSLKSASNYAYMDNTLKEVINOVEETGYNLLT 240
Db 234 SEMKNGYISADQYKAVNTPITDGLQSLKSASNYAYMDNTLKEVINOVEETGYNLLT 293
OY 241 TGMADVNTNDOEAKHLMIDYNTDEVAYAPDELAQVASTIVDVSNKVIQAOLGARHOSSN 300
Db 294 TGMADVNTNDOEAKHLMIDYNTDEVAYAPDELAQVASTIVDVSNKVIQAOLGARHOSSN 353
OY 301 VSFGLNQAVETNRDWSGTMKPTTDVAPALEYGYDSTATIVHDEPNYNGCTNTPVYNMNR 360
Db 354 VSFGLNQAVETNRDWSGTMKPTTDVAPALEYGYDSTATIVHDEPNYNGCTNTPVYNMNR 413
OY 361 GYFGNTTLOYALQOOSNNVAVETLNKVGILNRAKTFPLNGLGIDIPSIHYSAISSNTTESD 420
Db 414 AVFGNTTLOYALQOOSNNVAVETLNKVGILNRAKTFPLNGLGIDIPSIHYSAISSNTTESN 473
```

```
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTTAYMTMD 480
Db 474 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTTAYMTMD 533
OY 481 MKKTVLTGTGRNAYLAMPLOAGKGTGTSNYTDEIEINHIKTSOFVAPDELFACTYRKYSK 540
Db 534 MKKTVLTGTGRNAYLAMPLOAGKGTGTSNYTDEIEINHIKTSOFVAPDELFACTYRKYSK 593
OY 541 AVWTGYSNRLTPLYGNGLTVAAKVYRSMTYIYSEGSNPEDMNIPEGLYRNGEFVFNKGAR 600
Db 594 AVWTGYSNRLTPLYGNGLTVAAKVYRSMTYIYSEGSNPEDMNIPEGLYRNGEFVFNKGAR 653
OY 601 STWNSPAPQOPPESTESSSSSDSTSSSTPTSTNNSTTTNNNTQOOSNTTPOOONON 660
Db 654 PWTERTQOOSSTAESSSSDSTSSSTPTSTNNSTTTNNNTQOOSNTTPOOONON 713
OY 661 POPAP 666
Db 714 POPAP 719
```

```
RESULT 7
S28032
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 681) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28032
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010977
A:Accession: S28032
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-References: EMBL:X67866; NID:947407; PIDN:CAA48066.1; PID:947408
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 93.1%; Score 3245; DB 2; Length 719;
Best Local Similarity 92.0%; Pred. No. 3,2e-180;
Matches 613; Conservative 35; Mismatches 18; Indels 0; Gaps 0;
OY 1 KIYDNKNOLIADGSERRVNAQANDIPTDLVKAIVSIEDHRRFDHGIDIRILGAFLRN 60
Db 54 KIYDNKNOLIADGSERRVNAQANDIPTDLVKAIVSIEDHRRFDHGIDIRILGAFLRN 113
OY 61 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWMLAIQLEQKATKQETLTYIYINK 120
Db 114 LOSNSLOGGSTLTQOLIKLTYFSTSDQTSRKAQEWMLAIQLEQKATKQETLTYIYINK 173
OY 121 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRNLVL 180
Db 174 VYMSNGNYGMOTRAQNYKYKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRNLVL 233
OY 181 SEMKNGYISADQYKAVNTPITDGLQSLKSASNYAYMDNTLKEVINOVEETGYNLLT 240
Db 234 SEMKNGYISADQYKAVNTPITDGLQSLKSASNYAYMDNTLKEVINOVEETGYNLLT 293
OY 241 TGMADVNTNDOEAKHLMIDYNTDEVAYAPDELAQVASTIVDVSNKVIQAOLGARHOSSN 300
Db 294 TGMADVNTNDOEAKHLMIDYNTDEVAYAPDELAQVASTIVDVSNKVIQAOLGARHOSSN 353
OY 301 VSFGLNQAVETNRDWSGTMKPTTDVAPALEYGYDSTATIVHDEPNYNGCTNTPVYNMNR 360
Db 354 VSFGLNQAVETNRDWSGTMKPTTDVAPALEYGYDSTATIVHDEPNYNGCTNTPVYNMNR 413
OY 361 GYFGNTTLOYALQOOSNNVAVETLNKVGILNRAKTFPLNGLGIDIPSIHYSAISSNTTESD 420
Db 414 AVFGNTTLOYALQOOSNNVAVETLNKVGILNRAKTFPLNGLGIDIPSIHYSAISSNTTESN 473
OY 421 KRYGASSEKMAAAYAFANGGTYRKPMYIHKVYFSDGSEKESNVGTRAMKETTTAYMTMD 480
```


Db	474	KÖYGAŞEKMAAAYAAPANGGIYHKPPYINKVYFSDGSRKEPESDVGTBAMKETTAYMMTE	533
OY	481	MMKTVLTGTGTGNATLAMPQAGKTGTSNYDELENNHKTSQFAPAPPELFEGYTRKXSM	540
Db	534	MMKTVLALGTGTGATLPWLAQRKGTGTSNYDDELKHKTSQYVAPDPMFVGYTRKXSM	593
OY	541	AVMTGYSNRLPPLVONGSLTVAAKYVRSMTYTLSEGSNPEDMNIPBGLYRNGEVEFKNGAR	600
Db	594	AVMTGYSNRLPPIVODGGLVAAKYVRSMTIYLSBDTHPEDMTMPDGLFRNGEVEFKNGAR	653
OY	601	STWNSPAPQPPSTPSESSSSSDSSSTPSTPNNSTTTPPNNNTPOOSONTTPPOOON	660
Db	654	STWNSPAPQPPSTPSESSSSSDSSSTPSTPNNSTTTPPNNNTPOOSONTTPPOOON	713
OY	661	POPAPQ 666	
Db	714	POPAPQ 719	

RESULT 8.
S28031
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 456) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28031
R:Martin, C., Sibold, C., Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant strains of *Streptococcus pneumoniae*
A:Reference number: S28031; MUID:39010977
A:Accession: S28031
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67868; NID:947405; PIDN:CAA48068.1; PID:947406
;Superfamily: penicillin-binding protein 1B

[illegible]

```

QY 481 MWVLYVTGTRNNVYLAMLPOAGKTGTSNTDEIEKHITTSQFVAPDELFACYTRKYS 540
Db 534 MKKVLTYLTGTGRGATLPWLPQAGKTGTSNTDEEVENHINKTGVADEMFVGYTRKYS 593
QY 541 AVMTGYSNRLPLVPGONGLTVAAKYVYRSMMTYLSEGSNPEDMNIPEGLYRNGEYVFRNGAR 600
Db 594 AVMTGYSNRLPLIYGDGFLVAAKYVYRSMTLYLSDTHPEDMWTMPDGLFRNGEYVFRNGAR 653
QY 601 STWNSPAFOQPPSTBSSSSSSDSTSOSSSTTPTSTNNSTTTPNNNTQOSNTTPDOONON 660
Db 654 PIWTEPSTSOOSSTBSSSSSSDSTSOSSSTTPTSTNNSTTTPNNNTQOSNTTPDOONON 713
QY 661 POPAPQ 666
Db 714 POPAPQ 719

```

RESULT 9
S28033
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 670) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28033
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010957
A:Accession: S28033
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67867; NID:g47409; PIDN:CAA48067.1; PID:g47410
C:Superfamily: penicillin-binding protein 1B

Query Match	Similarity	93.0%	Score 3241	DB 2	Length 719
Best Local	612	Conservative	36	Mismatches	18
Indels	0	Gaps	0		
Pred. No. 5.5e-180					
Matches	612	Conservative	36	Mismatches	18
Indels	0	Gaps	0		
1	KTYDNKNQIADLGSERRNAOANDIPTDLVAIVASIEDHRRFDHRCITIRILGAFLRN	60			
54	KTYDNKNQIADLGSERRNAOANDIPTDLVAIVASIEDHRRFDHRCIDISIRILGAFLRN	113			
61	IQSNLSGGSTLTQOLIKLTFSTSTSDQTSIKRAQEAMLAIQLEQAKRQELITFYINK	120			
114	IQSNLSGGSTLTQOLIKLTFSTSTSDQTSIKRAQEAMLAIQLEQAKRQELITFYINK	173			
121	VYMSNGNVMQTAOANYCYKDLNNLSLPOLALLAGPQAPNOYDPSHPBAADRNLVL	180			
174	VYMSNGNVMQTAOANYCYKDLNNLSLPOLALLAGPQAPNOYDPSHPBAADRNLVL	233			
181	SMKMGNGVTSABOYEKAVNPTPTDGLQSLKASNPNYAVMDNYLKEVINQEEETGNLLT	240			
234	SMKMGNGVTSABOYEKAVNPTPTDGLQSLKASNPNYAVMDNYLKEVINQEEETGNLLT	293			
241	TGMDVYTNDOEAOKHLMIDIVYTDVEVAYPDDLEQVASTIVDVSNNKCVIAQLCARHQQSN	300			
294	TGMDVYTNDOEAOKHLMIDIVYTDVEVAYPDDLEQVASTIVDVSNNKCVIAQLCARHQQSN	353			
301	VSEFGINQAVETNRDWSITMKPTTDYAPALEYGVYSTATIVHDEPNYPTGNTPNVNMDR	360			
354	VSEFGINQAVETNRDWSITMKPTTDYAPALEYGVYSTATIVHDEPNYPTGNTPNVNMDR	413			
361	GTFGNITLQYALQOOSNNVAVETLNKVGILNRAKTFPLNGIGIDYPSLHYSMAISNTTESD	420			
414	AFYGNITLQYALQOOSNNVAVETLNKVGILNRAKTFPLNGIGIDYPSLHYSMAISNTTESN	473			
421	KYVGASSEKMAAAYAAAFANGGYFYKRPYTHKYVVSQSEKESNVGTRPMKETTAAVMMTD	480			
474	KYVGASSEKMAAAYAAAFANGGYTHKRPYTHKYVVSQSEKESNVGTRPMKETTAAVMMTE	533			
481	NMKTYLVYGTGNAYLAWLPQAGKTGTSNYTDEELINNHKTQSFVAPDELFAGYTFKRYSN	540			

```
||||| ||||| ||| || ||||| ||||| ||| : ||||| : ||||| |||||
534 MKKTVLAVGTGRCAYLPLMLAQAGKTGTSNYTDEIEKHITKNGYVAPDEMFYTRKYSM 593
Db
Oy 541 AWTGYSNRLTPLYVNGLTVAAKVYRSMTYISSEGNPEPDWNIPEGLYNGEYVFNKNGAR 600
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 594 AWTGYSNRLTPLYVNGLTVAAKVYRSMTYISSEGNPEPDWNIPEGLYNGEYVFNKNGAR 653
Oy 601 STMNSPAPQPPSTESSSSSDSTSSSTPTSTNNSTTTNNNTQOQNTTTPDOQON 660
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 654 STMNSPAPQPPSTESSSSSDSTSSSTPTSTNNSTTTNNNTQOQNTTTPDOQON 713
Oy 661 POPAP 666
||| |||||
Db 714 POPAP 719
```

```
RESULT 10
S28035
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 2039) (fragment)
C:Species: Streptococcus pneumoniae
A:Variety: strain 2039
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28035
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A>Title: Relatedness of penicillin-binding protein 1a genes from different clones of pen
A:Reference number: S28031; MUID:93010977
A:Accession: S28035
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67870; NID:g47413; PIDN:CAA48070.1; PID:g47414
A:Experimental source: strain 2039
C:Superfamily: penicillin-binding protein 1B
C:Keywords: antibiotic resistance; penicillin resistance
```

```
Query Match 91.6%; Score 3192; DB 2; Length 719;
Best Local Similarity 90.4%; Pred. No. 3.8e-177;
Matches 602; Conservative 41; Mismatches 23; Indels 0; Gaps 0;

Oy 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 60
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54 KIFDSKNELIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 113
Oy 61 LOSNSLQGGSTLTQOLIKLTFSTSDQTSRKAQEWALAIQLEKATKOEILTYINK 120
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 LOSNSLQGGSTLTQOLIKLTFSTSDQTSRKAQEWALAIQLEKATKOEILTYINK 173
Oy 121 VYMSNGNYGMQTPAQNYYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRLVL 180
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 174 VYMSNGNYGMQTPAQNYYGKDLRELSTPOLALLAGMPQAPNOYDPSHPEAADRRLVL 233
Oy 181 SEMKNGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOVEEETGNLIT 240
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234 SEMKNGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOVEEETGNLIT 293
Oy 241 TGMADVNTNDOEAKRLMDIYNTDEVYAPDDELQVASTIVVSNKVAIQAQARHOSN 300
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 294 TGMADVNTNDOEAKRLMDIYNTDEVYAPDDELQVASTIVVSNKVAIQAQARHOSN 353
Oy 301 VSEGINQAVETNRDMSITKPTIDYAPALEYGYVDSSTATIVHDEPNYGTPTPVYNNMR 360
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 354 VSEGINQAVETNRDMSITKPTIDYAPALEYGYVDSSTATIVHDEPNYGTPTPVYNNMR 413
Oy 361 GYFGNTTLOALAOQSNRYAVETLNKVGILNRAKTFPLNGIGIDYPSLHYSNALISSNTESD 420
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 414 GYFGNTTLOALAOQSNRYAVETLNKVGILNRAKTFPLNGIGIDYPSLHYSNALISSNTESN 473
Oy 421 KRYGASSEMMAAAYAAAFANGGTYRKPMYTHKVVFSGSEKESNVGTRAMKETATAMMD 480
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 474 KRYGASSEMMAAAYAAAFANGGTYRKPMYTHKVVFSGSEKESNVGTRAMKETATAMMD 533
474 KRYGASSEMMAAAYAAAFANGGTYRKPMYTHKVVFSGSEKESNVGTRAMKETATAMMD 533
```

```
Oy 481 MKKTVLVTGGRNAYLAWLPOAGKTGTSNYTDEIEKHITKNGYVAPDEMFYTRKYSM 540
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 534 MKKTVLADIGIGAYLPLMLAQAGKTGTSNYTDEIEKHITKNGYVAPDEMFYTRKYSM 593
Oy 541 AWTGYSNRLTPLYVNGLTVAAKVYRSMTYISSEGNPEPDWNIPEGLYNGEYVFNKNGAR 600
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 594 AWTGYSNRLTPLYVNGLTVAAKVYRSMTYISSEGNPEPDWNIPEGLYNGEYVFNKNGAR 653
Oy 601 STMNSPAPQPPSTESSSSSDSTSSSTPTSTNNSTTTNNNTQOQNTTTPDOQON 660
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 654 STMNSPAPQPPSTESSSSSDSTSSSTPTSTNNSTTTNNNTQOQNTTTPDOQON 713
Oy 661 POPAP 666
||| |||||
Db 714 POPAP 719
```

```
RESULT 11
S28036
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 8250) (fragment)
C:Species: Streptococcus pneumoniae
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28036
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A>Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MUID:93010977
A:Accession: S28036
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-608 <MAR>
A:Cross-references: EMBL:X67871; NID:g47415; PIDN:CAA48071.1; PID:g47416
C:Superfamily: penicillin-binding protein 1B
```

```
Query Match 80.9%; Score 2820; DB 2; Length 608;
Best Local Similarity 96.8%; Pred. No. 1e-195;
Matches 537; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

Oy 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 60
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 54 KIFDSKNELIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 113
Oy 61 LOSNSLQGGSTLTQOLIKLTFSTSDQTSRKAQEWALAIQLEKATKOEILTYINK 120
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 114 LOSNSLQGGSTLTQOLIKLTFSTSDQTSRKAQEWALAIQLEKATKOEILTYINK 173
Oy 121 VYMSNGNYGMQTPAQNYYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRLVL 180
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 174 VYMSNGNYGMQTPAQNYYGKDLNNLSLPOLALLAGMPQAPNOYDPSHPEAADRRLVL 233
Oy 181 SEMKNGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOVEEETGNLIT 240
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 234 SEMKNGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLTKEVINOVEEETGNLIT 293
Oy 241 TGMADVNTNDOEAKRLMDIYNTDEVYAPDDELQVASTIVVSNKVAIQAQARHOSN 300
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 294 TGMADVNTNDOEAKRLMDIYNTDEVYAPDDELQVASTIVVSNKVAIQAQARHOSN 353
Oy 301 VSEGINQAVETNRDMSITKPTIDYAPALEYGYVDSSTATIVHDEPNYGTPTPVYNNMR 360
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 354 VSEGINQAVETNRDMSITKPTIDYAPALEYGYVDSSTATIVHDEPNYGTPTPVYNNMR 413
Oy 361 GYFGNTTLOALAOQSNRYAVETLNKVGILNRAKTFPLNGIGIDYPSLHYSNALISSNTESD 420
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 414 GYFGNTTLOALAOQSNRYAVETLNKVGILNRAKTFPLNGIGIDYPSLHYSNALISSNTESD 473
Oy 421 KRYGASSEMMAAAYAAAFANGGTYRKPMYTHKVVFSGSEKESNVGTRAMKETATAMMD 480
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 474 KRYGASSEMMAAAYAAAFANGGTYRKPMYTHKVVFSGSEKESNVGTRAMKETATAMMD 533
481 MKKTVLVTGGRNAYLAWLPOAGKTGTSNYTDEIEKHITKNGYVAPDEMFYTRKYSM 540
||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Db 709 AEDKKTEEEKKEEAK 727

Search completed: June 13, 2002, 08:41:08
Job time: 34 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:41:09 ; Search time 34.23 Seconds

(without alignments)
3365.896 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484
Sequence: 1 KIYDNKNQIADLGSERRVN.....TQSNMTPDQNGNQPAQP 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhcc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	719	2	Q9REU0 streptococc
2	3468	99.5	719	2	Q9REU9 streptococc
3	3403	97.7	719	2	Q9REU3 streptococc
4	3329	95.6	719	2	Q9WVW0 streptococc
5	3314	95.1	719	2	Q9REU8 streptococc
6	3299	94.7	719	2	Q9WVW1 streptococc
7	3274	94.0	719	2	Q9REU4 streptococc
8	3247	93.2	719	2	Q9REU4 streptococc
9	3246	93.2	719	2	Q9REU6 streptococc
10	3245	93.2	719	2	Q54949 streptococc
11	3243	93.1	719	2	Q54947 streptococc
12	3243	93.0	719	2	Q54946 streptococc
13	3241	92.7	719	2	Q54948 streptococc
14	3231	91.6	719	2	Q9REU5 streptococc
15	3192	90.0	719	2	Q54950 streptococc
16	3134	90.0	718	2	Q9REU7 streptococc

17	2820	80.9	608	2	Q54951 streptococc
18	2791	80.1	728	2	Q9F2G7 streptococc
19	2696	77.4	727	2	Q9F2G4 streptococc
20	2432	69.8	527	2	Q70037 streptococc
21	2091	60.0	398	2	Q9R7M6 streptococc
22	2088	59.9	398	2	Q87102 streptococc
23	2037.5	58.5	721	16	Q99YX1 streptococc
24	2029	58.2	398	2	Q9R7M5 streptococc
25	2015	57.8	398	2	Q87108 streptococc
26	2011	57.7	398	2	Q87105 streptococc
27	1936	55.6	398	2	Q9R7M7 streptococc
28	1926	55.3	398	2	Q87103 streptococc
29	1921	55.1	398	2	Q9R7M8 streptococc
30	1911	54.9	793	2	Q9EXM9 streptococc
31	1881	54.0	398	2	Q87107 streptococc
32	1876	53.8	398	2	Q87104 streptococc
33	1757	50.4	400	2	Q87106 streptococc
34	1751	50.3	778	2	Q9EXN1 streptococc
35	1657.5	47.6	664	16	Q9C123 streptococc
36	1592	45.7	310	2	Q52741 streptococc
37	1583	45.4	310	2	Q52743 streptococc
38	1573	45.1	310	2	Q52739 streptococc
39	1571	45.1	310	2	Q52737 streptococc
40	1567	45.0	310	2	Q52738 streptococc
41	1564	44.9	310	2	Q52740 streptococc
42	1558	44.7	310	2	Q52742 streptococc
43	1454	41.7	310	2	Q93N72 streptococc
44	1449	41.6	310	2	Q52744 streptococc
45	1201	34.5	826	16	Q92AB6 streptococc

ALIGNMENTS

RESULT 1
Q9REU0 PRELIMINARY; PRT; 719 AA.
AC Q9REU0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRATN-SP 1261;
RX MEDLINE-21432820; PubMed-11549185;
RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
RT penicillin-resistant clinical isolates of Streptococcus pneumoniae
RT serotype 23F from the nasopharyngeal flora of children.";
RI J. Med. Microbiol. 50:828-832(2001).
DR EMBL: AF210745; AAF17255.1;
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
SO SEQUENCE 719 AA; 79758 MW; 5BD397EB3BA63AA6 CRC64;

Query Match 100.0%; Score 3484; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 2, 1e-181;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRFPDHRGIDTIRILGAFARN 60
Db 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRFPDHRGIDTIRILGAFARN 113

```
QY 61 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKAQEAMLAIQLEOKATKQELITTYINK 120
    |||||
Db 114 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKAQEAMLAIQLEOKATKQELITTYINK 173
QY 121 VYMSNGNYGMQTAQNYKYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRNLV 180
    |||||
Db 174 VYMSNGNYGMQTAQNYKYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRNLV 233
QY 181 SEMKNGYISAEQYERKAVNTPITDGLQSLKSASNPAYMDNYLKEYINQVEETGYNLLT 240
    |||||
Db 234 SEMKNGYISAEQYERKAVNTPITDGLQSLKSASNPAYMDNYLKEYINQVEETGYNLLT 293
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAAPDELOVASTIVDYSNGKVIAQLGARHQS 300
    |||||
Db 294 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAAPDELOVASTIVDYSNGKVIAQLGARHQS 353
QY 301 VSFGINQAVETNRDMSGTYKPTITDVAFALEYGYDSTATIVHDEPNYNGTNPVYNM 360
    |||||
Db 354 VSFGINQAVETNRDMSGTYKPTITDVAFALEYGYDSTATIVHDEPNYNGTNPVYNM 413
QY 361 GYFGNITLQYALQOQRNVPVAVETLNKVGILNRAKTFPLNGIDIPSIHYSNAISSNT 420
    |||||
Db 414 GYFGNITLQYALQOQRNVPVAVETLNKVGILNRAKTFPLNGIDIPSIHYSNAISSNT 473
QY 421 KRYGASSEKMAAAYAFANGGTYKPMYIHKVYFSDGSEKEFSNVGTRAKETAYAMTD 480
    |||||
Db 474 KRYGASSEKMAAAYAFANGGTYKPMYIHKVYFSDGSEKEFSNVGTRAKETAYAMTD 533
QY 481 MKTIVTYTGGRNAYLAMLPOAGKTGTSNTDEIENHIKTSQFVAPDELFACTYKRY 540
    |||||
Db 534 MKTIVTYTGGRNAYLAMLPOAGKTGTSNTDEIENHIKTSQFVAPDELFACTYKRY 593
QY 541 AVWTGYSNRLTPLYNGGLTYAAKVRYSMTYLSGGSNPEDMNIPEGLYRNGEYVFN 600
    |||||
Db 594 AVWTGYSNRLTPLYNGGLTYAAKVRYSMTYLSGGSNPEDMNIPEGLYRNGEYVFN 653
QY 601 STWNSPAPQOPSTESSSSSDSSTSPSTNNSTTTNNNTQOQNTTPDOONON 660
    |||||
Db 654 STWNSPAPQOPSTESSSSSDSSTSPSTNNSTTTNNNTQOQNTTPDOONON 713
QY 661 POPAP 666
    |||||
Db 714 POPAP 719

RESULT 2
O9RET9 PRELIMINARY: PRT: 719 AA.
AC O9RET9;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP 1513;
RX MEDLINE=21432820; PubMed=11549185;
RA Feroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
RT penicillin-resistant clinical isolates of Streptococcus pneumoniae
RT serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00903; Transglycosyl; 1.
DR Pfam: PF00903; Transpeptidase; 1.
DR Pdbom: PD001895; Transglycosyl; 1.
```

```
SQ SEQUENCE 719 AA: 79830 MW: 5f6776b8DEAE3840 CRC64:
Query Match 99.5%; Score 3468; DB 2; Length 719;
Best Local Similarity 99.5%; Pred. No. 1.6e-180;
Matches 663; Conservative 2; Mismatches 1; Indels 0; Gaps 0:
QY 1 KIDNKNOLIADIGSERRRVNAQANDIPTDLVKAIVSIEDHFRFDRHGDITRIILGAFLRN 60
    |||||
Db 54 KIDNKNOLIADIGSERRRVNAQANDIPTDLVKAIVSIEDHFRFDRHGDITRIILGAFLRN 113
QY 61 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKAQEAMLAIQLEOKATKQELITTYINK 120
    |||||
Db 114 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKRKAQEAMLAIQLEOKATKQELITTYINK 173
QY 121 VYMSNGNYGMQTAQNYKYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRNLV 180
    |||||
Db 174 VYMSNGNYGMQTAQNYKYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRRNLV 233
QY 181 SEMKNGYISAEQYERKAVNTPITDGLQSLKSASNPAYMDNYLKEYINQVEETGYNLLT 240
    |||||
Db 234 SEMKNGYISAEQYERKAVNTPITDGLQSLKSASNPAYMDNYLKEYINQVEETGYNLLT 293
QY 241 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAAPDELOVASTIVDYSNGKVIAQLGARHQS 300
    |||||
Db 294 TGMDEVYTNVDOEAQKHLMDIYNTDEVYAAPDELOVASTIVDYSNGKVIAQLGARHQS 353
QY 301 VSFGINQAVETNRDMSGTYKPTITDVAFALEYGYDSTATIVHDEPNYNGTNPVYNM 360
    |||||
Db 354 VSFGINQAVETNRDMSGTYKPTITDVAFALEYGYDSTATIVHDEPNYNGTNPVYNM 413
QY 361 GYFGNITLQYALQOQRNVPVAVETLNKVGILNRAKTFPLNGIDIPSIHYSNAISSNT 420
    |||||
Db 414 GYFGNITLQYALQOQRNVPVAVETLNKVGILNRAKTFPLNGIDIPSIHYSNAISSNT 473
QY 421 KRYGASSEKMAAAYAFANGGTYKPMYIHKVYFSDGSEKEFSNVGTRAKETAYAMTD 480
    |||||
Db 474 KRYGASSEKMAAAYAFANGGTYKPMYIHKVYFSDGSEKEFSNVGTRAKETAYAMTD 533
QY 481 MKTIVTYTGGRNAYLAMLPOAGKTGTSNTDEIENHIKTSQFVAPDELFACTYKRY 540
    |||||
Db 534 MKTIVTYTGGRNAYLAMLPOAGKTGTSNTDEIENHIKTSQFVAPDELFACTYKRY 593
QY 541 AVWTGYSNRLTPLYNGGLTYAAKVRYSMTYLSGGSNPEDMNIPEGLYRNGEYVFN 600
    |||||
Db 594 AVWTGYSNRLTPLYNGGLTYAAKVRYSMTYLSGGSNPEDMNIPEGLYRNGEYVFN 653
QY 601 STWNSPAPQOPSTESSSSSDSSTSPSTNNSTTTNNNTQOQNTTPDOONON 660
    |||||
Db 654 STWNSPAPQOPSTESSSSSDSSTSPSTNNSTTTNNNTQOQNTTPDOONON 713
QY 661 POPAP 666
    |||||
Db 714 POPAP 719

RESULT 3
O9RET3 PRELIMINARY: PRT: 719 AA.
AC O9RET3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM 4200;
RX MEDLINE=21432820; PubMed=11549185;
```


RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1a, 2b and 2x amongst
 penicillin-resistant clinical isolates of *Streptococcus pneumoniae*
 serotype 23F from the nasopharyngeal flora of children.";
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL: AF210752; AAF17262.1; -.
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl. 1.
 DR Pfam: PF00905; Transpeptidase. 1.
 DR Prodom: PD001895; Transglycosyl. 1.
 SQ SEQUENCE 719 AA; 79781 MW; DB5993F6BD47F72D CRC64;

Query Match 97.7%; Score 3403; DB 2; Length 719;
 Best Local Similarity 97.3%; Pred. No. 5.4e-177;
 Matches 648; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSEKRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
 DB 54 KIYDNKNLIDLGSEKRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
 QY 61 LOSNSLOGGSLTLOOLIKLTYFSTSDQTSRKAEAMLAIOLEOKATKOEILTYINK 120
 DB 114 LOSNSLOGGSLTLOOLIKLTYFSTSDQTSRKAEAMLAIOLEOKATKOEILTYINK 173
 QY 121 VYMSNGNGMOTAAONNYGKDLNNLSLPOLALLGMPAPAPNOYDPSHEAODRRNLVL 180
 DB 174 VYMSNGNGMOTAAONNYGKDLNNLSLPOLALLGMPAPAPNOYDPSHEAODRRNLVL 233
 QY 181 SEMKNOGYISAEOYKAVNPTITDGLSLKSASNPAYMDNYLKEVINOVEETGYNLLT 240
 DB 234 SEMKNOGYISAEOYKAVNPTITDGLSLKSASNPAYMDNYLKEVINOVEETGYNLLT 293
 QY 241 TGMVYTNVDEAOKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKYIAOLGARHOSN 300
 DB 294 TGMVYTNVDEAOKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKYIAOLGARHOSN 353
 QY 301 VSFGINOVAVERNRMGSMKPTITDAPALEGYVDSTATIVHDEPYNPGTNPYYNMDR 360
 DB 354 VSFGINOVAVERNRMGSMKPTITDAPALEGYVDSTATIVHDEPYNPGTNPYYNMDR 413
 QY 361 GYFNGITLOVALOOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHNSNAISSWTESD 420
 DB 414 GYFNGITLOVALOOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHNSNAISSWTESD 473
 QY 421 KKYGASSEKMAAAYAAAFANGSTYKPMYIHKVYVSDGSEKEFSNVGTRAMKETTAYMMD 480
 DB 474 KKYGASSEKMAAAYAAAFANGSTYKPMYIHKVYVSDGSEKEFSNVGTRAMKETTAYMMD 533
 QY 481 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDELENIHKTQSOFAPADELFAGYTRKYSM 540
 DB 534 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDELENIHKTQSOFAPADELFAGYTRKYSM 593
 QY 541 AVMTGYSNRLTPLYGNGLTVAAYKYSMMTYLSGSPNEDNNIPGGLYRNGEFPVKNGAR 600
 DB 594 AVMTGYSNRLTPLYGNGLTVAAYKYSMMTYLSGSPNEDNNIPGGLYRNGEFPVKNGAR 653
 QY 601 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOOSNTTTPDOQN 660
 DB 654 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOOSNTTTPDOQN 713
 QY 661 POPAP 666
 DB 714 POPAP 719

RESULT 4
 Q9WVWO PRELIMINARY; PRT: 719 AA.
 AC Q9WVWO;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPLA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Filicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 ON NCBI_TaxID=1313;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PO-342, PO-273, AND PO-341;
 RA Coffey T.O., Daniels M., Enright M.C., Spratt B.G.;
 RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
 clone of *Streptococcus pneumoniae* arose by large recombinational
 RT replacements of the cpsA-pbpA region.";
 RL Microbiology 0:0-0(1999)
 DR EMBL: AF139886; AAD43069.1; -.
 DR EMBL: AF139884; AAD43067.1; -.
 DR EMBL: AF139885; AAD43068.1; -.
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl. 1.
 DR Pfam: PF00905; Transpeptidase. 1.
 DR Prodom: PD001895; Transglycosyl. 1.
 SQ SEQUENCE 719 AA; 79622 MW; 7FB6A75EA8FF8B3 CRC64;

Query Match 95.6%; Score 3329; DB 2; Length 719;
 Best Local Similarity 94.4%; Pred. No. 5.6e-173;
 Matches 629; Conservative 25; Mismatches 12; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSEKRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
 DB 54 KIYDNKNLIDLGSEKRVNAQANDIPDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
 QY 61 LOSNSLOGGSLTLOOLIKLTYFSTSDQTSRKAEAMLAIOLEOKATKOEILTYINK 120
 DB 114 LOSNSLOGGSLTLOOLIKLTYFSTSDQTSRKAEAMLAIOLEOKATKOEILTYINK 173
 QY 121 VYMSNGNGMOTAAONNYGKDLNNLSLPOLALLGMPAPAPNOYDPSHEAODRRNLVL 180
 DB 174 VYMSNGNGMOTAAONNYGKDLNNLSLPOLALLGMPAPAPNOYDPSHEAODRRNLVL 233
 QY 181 SEMKNOGYISAEOYKAVNPTITDGLSLKSASNPAYMDNYLKEVINOVEETGYNLLT 240
 DB 234 SEMKNOGYISAEOYKAVNPTITDGLSLKSASNPAYMDNYLKEVINOVEETGYNLLT 293
 QY 241 TGMVYTNVDEAOKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKYIAOLGARHOSN 300
 DB 294 TGMVYTNVDEAOKHLMIDYNTDEYVAYPDDELQVASTIVDVNSGKYIAOLGARHOSN 353
 QY 301 VSFGINOVAVERNRMGSMKPTITDAPALEGYVDSTATIVHDEPYNPGTNPYYNMDR 360
 DB 354 VSFGINOVAVERNRMGSMKPTITDAPALEGYVDSTATIVHDEPYNPGTNPYYNMDR 413
 QY 361 GYFNGITLOVALOOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHNSNAISSWTESD 420
 DB 414 GYFNGITLOVALOOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHNSNAISSWTESD 473
 QY 421 KKYGASSEKMAAAYAAAFANGSTYKPMYIHKVYVSDGSEKEFSNVGTRAMKETTAYMMD 480
 DB 474 KKYGASSEKMAAAYAAAFANGSTYKPMYIHKVYVSDGSEKEFSNVGTRAMKETTAYMMD 533
 QY 481 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDELENIHKTQSOFAPADELFAGYTRKYSM 540
 DB 534 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDELENIHKTQSOFAPADELFAGYTRKYSM 593
 QY 541 AVMTGYSNRLTPLYGNGLTVAAYKYSMMTYLSGSPNEDNNIPGGLYRNGEFPVKNGAR 600
 DB 594 AVMTGYSNRLTPLYGNGLTVAAYKYSMMTYLSGSPNEDNNIPGGLYRNGEFPVKNGAR 653
 QY 601 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOOSNTTTPDOQN 660
 DB 654 STWNSPAPQPPSTRESSSSSDSSTSPSTNNSTTNPNNNTQOOSNTTTPDOQN 713

OY 661 POPAP 666
 |||||
 Db 714 POPAP 719

RESULT 5

O9RET8 PRELIMINARY: PRT: 719 AA.

AC O9RET8: 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP 1465;
 RX MEDLINE=21432820; PubMed=11549185;
 RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children."
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL, AF210747; AAF17257.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR Prodom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79662 MW; A7B96EA15434A805 CRC64;

Query Match 95.1%; Score 3314; DB 2; Length 719;
 Best Local Similarity 94.1%; Pred. No. 3.7e-172;
 Matches 627; Conservative 27; Mismatches 12; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFPDHGIPTIRILGAFLRN 60
 Db 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFPDHGIPTIRILGAFLRN 113
 OY 61 LQNSLSGGSTLTQOLIKLTFYFSTSDQTSRKAQEWLAIOLEKATKOEILTYINK 120
 Db 114 LQNSLSGGSTLTQOLIKLTFYFSTSDQTSRKAQEWLAIOLEKATKOEILTYINK 173
 OY 121 VYMSNGNYGMQTAQNYGYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPBAADRNLVY 180
 Db 174 VYMSNGNYGMQTAQNYGYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPBAADRNLVY 233
 OY 181 SEMKNGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNYIKVINYOVEETGNLLT 240
 Db 234 SEMKGGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNYIKVINYOVEETGNLLT 293
 OY 241 TGMADYTYNDQEAOKHLMIDYNTDEVVAYPDELOVASTIVDSNGKVAIOLGARQSSN 300
 Db 294 TGMADYTYNDQEAOKHLMIDYNTDEVVAYPDELOVASTIVDSNGKVAIOLGARQSSN 353
 OY 301 VSEGINQAVETNRDMGSTMKPTTDVAPALEYGVYDSTATIVHDEPNYEGTNTPVYNNDR 360
 Db 354 VSEGINQAVETNRDMGSTMKPTTDVAPALEYGVYDSTATIVHDEPNYEGTNTPVYNNDR 413
 OY 361 GYFGNTTLOYALQOSRNPVAVETLKNKVGILNRAKTEPLNGIGIDYPSHYNAISSNTTESD 420
 Db 414 AFGNTTLOYALQOSRNPVAVETLKNKVGILNRAKTEPLNGIGIDYPSHYNAISSNTTES 473
 OY 421 KYVGASSEKMAAAYAFANAGTYRKPMYTHKVEYSDGSEKESPNVETRAMKETATAMMD 480
 Db 474 KYVGASSEKMAAAYAFANAGTYRKPMYTHKVEYSDGSEKESPNVETRAMKETATAMMD 533

OY 481 MKKTVLTGCTGRNAYLAWLPQAGKTGTSNYTDEIEINHIKTSQFVAPDELFAGYTRKYSM 540
 Db 534 MKKTVLTGCTGRNAYLAWLPQAGKTGTSNYTDEIEINHIKTSQFVAPDELFAGYTRKYSM 593
 OY 541 AVWTGYSNRLTPLVGNGLVAAKAVYRSMMTYLSEGSNPEDMNIPECLYRNGEYFNKAR 600
 Db 594 AVWTGYSNRLTPLVGNGLVAAKAVYRSMMTYLSEGSNPEDMNIPECLYRNGEYFNKAR 653
 OY 601 STMNSPAPQOPSTESSSSSDSSTSSSTPTSTNNSTTTNNNTQOONTPDOONON 660
 Db 654 STMNSPAPQOPSTESSSSSDSSTSSSTPTSTNNSTTTNNNTQOONTPDOONON 713
 OY 661 POPAP 666
 |||||
 Db 714 POPAP 719

RESULT 6

O9WM11 PRELIMINARY: PRT: 719 AA.

AC O9WM11: 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-URU-E159, M134, URU-E135, AND URU-E157;
 RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
 RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V replacements of the cpsA-pdp1a region."
 RL Microbiology 145:0-0(1999).
 DR EMBL, AF138890; AAD43073.1; -;
 DR EMBL, AF138887; AAD43070.1; -;
 DR EMBL, AF138886; AAD43071.1; -;
 DR EMBL, AF138889; AAD43072.1; -;
 DR EMBL, AF159448; AAD48430.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR Prodom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79701 MW; 3EDC8A21C94DAC87 CRC64;

Query Match 94.7%; Score 3299; DB 2; Length 719;
 Best Local Similarity 93.4%; Pred. No. 2.4e-171;
 Matches 622; Conservative 30; Mismatches 14; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFPDHGIPTIRILGAFLRN 60
 Db 54 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFPDHGIPTIRILGAFLRN 113
 OY 61 LQNSLSGGSTLTQOLIKLTFYFSTSDQTSRKAQEWLAIOLEKATKOEILTYINK 120
 Db 114 LQNSLSGGSTLTQOLIKLTFYFSTSDQTSRKAQEWLAIOLEKATKOEILTYINK 173
 OY 121 VYMSNGNYGMQTAQNYGYKDLNNLSLPOLALLAGMPQAPNOYDPYSHPBAADRNLVY 180
 |||||

```
Db 174 VYMSGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAAQDRRLVL 233
QY 181 SEMKNOGYISAQYKAVNTPTDGLQSLKSASNPAYMDNLTKEVINQVEETGYNLT 240
Db 234 SEMKNOGYISAQYKAVNTPTDGLQSLKSASNPAYMDNLTKEVINQVEETGYNLT 293
QY 241 TGMVYTVNDQEAQKHLMDIYNTEYVAYPDELOVASTIVDSNGKVIAGLGAHQSNN 300
Db 294 TGMVYTVNDQEAQKHLMDIYNTEYVAYPDELOVASTIVDSNGKVIAGLGAHQSNN 353
QY 301 VSFGINQAVETNRDWSMTKPTTDYAPALEYGVYSTATYVHDEYNTPGTPTPYNMNR 360
Db 354 VSFGINQAVETNRDWSMTKPTTDYAPALEYGVYSTATYVHDEYNTPGTPTPYNMNR 413
QY 361 GFYGNITLOVALQOQRNPAVETTLNKVGLNRAKTFPLNGLGIDYPSIHSNAISSNTESD 420
Db 414 AYFGNITLOVALQOQRNPAVETTLNKVGLNRAKTFPLNGLGIDYPSIHSNAISSNTESN 473
QY 421 KRYGASSEKMAAAYAAAFANGGTYRKPMYIHKVFSGSEKSEFNSVGTBRAMKETTAYMMDT 480
Db 474 KRYGASSEKMAAAYAAAFANGGTYRKPMYIHKVFSGSEKSEFNSVGTBRAMKETTAYMMDT 533
QY 481 MKKTYLTYGTGRNAYLAPLPAQAGTGTSTYDELEHNHITSQFAPADELFGATRRKISM 540
Db 534 MKKTYLTYGTGRNAYLAPLPAQAGTGTSTYDELEHNHITSQFAPADELFGATRRKISM 593
QY 541 AVMTGYSNRLTPVNGGLTVAAKYRSMNTYLSGNSPEDNNIPGLYRNGEFVFKNGAR 600
Db 594 AVMTGYSNRLTPVNGGLTVAAKYRSMNTYLSGNSPEDNNIPGLYRNGEFVFKNGAR 653
QY 601 STWNSPAPQOPPESTESSSSSDSSSTPSTNNSTTTPNNNTQOOSNTTTPDOQN 660
Db 654 STWNSPAPQOPPESTESSSSSDSSSTPSTNNSTTTPNNNTQOOSNTTTPDOQN 713
QY 661 POPAP 666
Db 714 POPAP 719

RESULT 7
Q9RET4 PRELIMINARY; PRT; 719 AA.
AC Q9RET4;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP 22861;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferroni A., Berche P.;
RT "Alterations to penicillin-binding proteins 1a, 2b and 2x amongst
penicillin-resistant clinical isolates of Streptococcus pneumoniae
serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL; AF210751; AAF17261.1; -
DR InterPro; IPR001264; Transglycosyl.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00905; Transglycosyl. 1.
DR Pfam; PF00905; Transpeptidase.
DR PRODOM; PD001895; Transglycosyl. 1.
DR SEQUENCE 719 AA; 79567 MW; 7B6EDCCBB8CF286 CRC64;
```

Query Match 94.0%; Score 3274; DB 2; Length 719;
Best Local Similarity 92.9%; Pred. No. 5.5e-170;
Matches 619; Conservative 31; Mismatches 16; Indels 0; Gaps 0;

```
QY 1 KIYDNKNQIADLGSERRVNAQNDIPTDLVKAIVSIEDHREDFHGIIDTTRILGAFLRN 60
Db 54 KIYDNKNQIADLGSERRVNAQNDIPTDLVKAIVSIEDHREDFHGRIGDSIRILGAFLRN 113
QY 61 LOSNLSGGSTLTQOLLKLTFFSTSTSDQTSRKAQEMALAIQLQKATKQELITYYINK 120
Db 114 LOSNLSGGSTLTQOLLKLTFFSTSTSDQTSRKAQEMALAIQLQKATKQELITYYINK 173
QY 121 VYMSGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAAQDRRLVL 180
Db 174 VYMSGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAAQDRRLVL 233
QY 181 SEMKNOGYISAQYKAVNTPTDGLQSLKSASNPAYMDNLTKEVINQVEETGYNLT 240
Db 234 SEMKNOGYISAQYKAVNTPTDGLQSLKSASNPAYMDNLTKEVINQVEETGYNLT 293
QY 241 TGMVYTVNDQEAQKHLMDIYNTEYVAYPDELOVASTIVDSNGKVIAGLGAHQSNN 300
Db 294 TGMVYTVNDQEAQKHLMDIYNTEYVAYPDELOVASTIVDSNGKVIAGLGAHQSNN 353
QY 301 VSFGINQAVETNRDWSMTKPTTDYAPALEYGVYSTATYVHDEYNTPGTPTPYNMNR 360
Db 354 VSFGINQAVETNRDWSMTKPTTDYAPALEYGVYSTATYVHDEYNTPGTPTPYNMNR 413
QY 361 GFYGNITLOVALQOQRNPAVETTLNKVGLNRAKTFPLNGLGIDYPSIHSNAISSNTESD 420
Db 414 AYFGNITLOVALQOQRNPAVETTLNKVGLNRAKTFPLNGLGIDYPSIHSNAISSNTESN 473
QY 421 KRYGASSEKMAAAYAAAFANGGTYRKPMYIHKVFSGSEKSEFNSVGTBRAMKETTAYMMDT 480
Db 474 KRYGASSEKMAAAYAAAFANGGTYRKPMYIHKVFSGSEKSEFNSVGTBRAMKETTAYMMDT 533
QY 481 MKKTYLTYGTGRNAYLAPLPAQAGTGTSTYDELEHNHITSQFAPADELFGATRRKISM 540
Db 534 MKKTYLTYGTGRNAYLAPLPAQAGTGTSTYDELEHNHITSQFAPADELFGATRRKISM 593
QY 541 AVMTGYSNRLTPVNGGLTVAAKYRSMNTYLSGNSPEDNNIPGLYRNGEFVFKNGAR 600
Db 594 AVMTGYSNRLTPVNGGLTVAAKYRSMNTYLSGNSPEDNNIPGLYRNGEFVFKNGAR 653
QY 601 STWNSPAPQOPPESTESSSSSDSSSTPSTNNSTTTPNNNTQOOSNTTTPDOQN 660
Db 654 STWNSPAPQOPPESTESSSSSDSSSTPSTNNSTTTPNNNTQOOSNTTTPDOQN 713
QY 661 POPAP 666
Db 714 POPAP 719

RESULT 8
Q57114 PRELIMINARY; PRT; 719 AA.
AC Q57114;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C5111 (CEPHALOSPORIN-RESISTANT CLINICAL ISOLATE);
RX MEDLINE=96012191; PubMed=7574521;
RA Coffey T.J., Daniels M., McDougal L.K., Dowson C.G., Tenover F.C.,
Spratt B.G.;
RT "Genetic analysis of clinical isolates of Streptococcus pneumoniae
with high-level resistance to expanded-spectrum cephalosporins.";
RL Antimicrob. Agents Chemother. 39:1306-1313(1995).
DR EMBL; Z49095; CAA88918.1; -
```


RESULT	10		
ID	054949	PRELIMINARY;	PRT; 719 AA.
AC	054949;		
DT	01-NOV-1996 (TRIMBLrel. 01, Created)		
DT	01-NOV-1996 (TRIMBLrel. 01, Last sequence update)		
DT	01-JUN-2001 (TRIMBLrel. 17, Last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).		
GN	PONA.		
OS	Streptococcus pneumoniae.		
OC	Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
CC	Streptococcus.		
OX	NCBI_TaxID=1313;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=56742;		
RX	MEDLINE=93010977; PubMed=1396576;		
RA	Martin G., Sibold C., Hakenbeck R.;		
RT	"Relationship of penicillin-binding protein 1a genes from different		
RT	clones of penicillin-resistant Streptococcus pneumoniae isolated in		
RT	South Africa and Spain.";		
RL	EMBO J. 11:3831-3836(1992).		
RL	EMBL; X67869; CAA48069.1; -		
DR	InterPro; IPR001264; Transglycosyl.		
DR	InterPro; IPR001460; Transpeptidase.		
DR	Pfam; PF00902; Transglycosyl; 1.		
DR	Pfam; PF00905; Transpeptidase; 1.		
DR	ProDom; PD001895; Transglycosyl; 1.		
FT	NON_TER	719	719
SEQUENCE	719 AA; 79750 MW; 489B6C6EC0CB861CD CRC64;		

Query Match	93.28	Score 3246	DB 2	Length 719
Best Local Similarly	92.08	Pred. 1.8e-168		
Matches 613, Conservative	31	Mismatches 162	Indels 0	Gaps 0

Qy	1	KIIOKKNOIADJGSESRVNAQANDIPDILYKALVSTIEDHREFPHRIBOTIRILGALRN	60
Dd	54	KIYDKNOLADJGSESRVNAQANDIPDILYKALVSTIEDHREFPHRIBOTIRILGALRN	11.3
Qy	61	LOSNSLOGSITLTOOLIKLITFTSTSPQOTSRAQOEMALATOLEOKATKOELLTYINK	120
Dd	114	LOSNSLOGSITLTOOLIKLITFTSTSDQOTSRAQOEMALATOLEOKATKOELLTYINK	17.3
Qy	121	YVMSNGNYGMOPTAAQONYGYKDLNLSLPTOLALLAGMPQAPNOYDPSHPEAODRRMLYL	180
Dd	174	YVMSNGNYGMOPTAAQONYGYKDLNLSLPTOLALLAGMPQAPNOYDPSHPEAODRRMLYL	23.3
Qy	181	SEMKNOGITSAEJOYEKAVNPIITDGLQSLKSASNYPAYMONYKEVINOVEEETGYMLLT	24.0
Dd	234	SEMKNOGITSAEJOYEKAVNPIITDGLQSLKSASNYPAYMONYKEVINOVEEETGYMLLT	29.3
Qy	241	TGMQVYTNVDOEAKHLMIDYINPDEYVAYRPDELOVASTIYDVSNGKVIHQOLGARHSSN	30.0
Dd	294	TGMQVYTNVDOEAKHLMIDYINPDEYVAYRPDELOVASTIYDVSNGKVIHQOLGARHSSN	35.3
Qy	301	VSFQINQAVETNRBMGSTMKPITDYAPALEEGYVDSTATYIHDEBPYNYPGTNPVYWMDR	36.6
Dd	354	VSFQINQAVETNRBMGSAMKPIITDYAPALEEGYVDSTATYIHDEBPYNYPGTNPVYWMDR	41.3
Qy	361	GYFGNITQYALQOOSRNVPAVETLNKVGILNRAKFTLNGIGIDPSPISYHSAISNTTESD	42.0
Dd	414	AYFGNITQYALQOOSRNVPAVETLNKVGILNRAKFTLNGIGIDPSPISYHSAISNTTESN	47.3
Qy	421	KKYAGSSSEKMAAAVAAFANGSTYKKPMYIHKVYFSDSSSEKFEFNVGRAKETTAYMMTD	48.0
Dd	474	KOYAGSSSEKMAAAVAAFANGSTYKKPMYIHKVYFSDSSSEKFEFNVGRAKETTAYMMTD	53.3
Qy	481	MMKYVLYGTGRNRYLAWLPAGKGTGNSNTDEIEKHITISOPVADDELFCAGYTRYSM	54.0
Dd	534	MMKYVLYGTGRGAYLPLPLPAGKGTGNSNTTDEVEVHNKIKTGYVADDEMFVGYTRYSM	59.3
Qy	541	AVWTGYSNRLPLVNGNGLYVAAKYRSGMNTYVLSGCSNPDEDMNIPDEGLYRGEVFEFKGAR	60.0

[illegible]

RESULT	11		
054947			
ID	054947	PRELIMINARY:	PRT; 719 AA.
AC	054947;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).		
GN	PONA.		
OS	Streptococcus pneumoniae.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;		
OX	Streptococcus.		
OX	NCBI_TaxID=1313;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=681;		
RX	MEDLINE=93010977; Pubmed=1396576;		
RA	Martin C., Sibold C., Hakenbeck R.;		
RT	"Relatedness of penicillin-binding protein 1a genes from different		
RT	clones of penicillin-resistant Streptococcus pneumoniae isolated in		
RT	South Africa and Spain.";		
RL	EMBO J. 11:3831-3836(1992).		
EMBL	X67866; CAA48066.1.		
DR	InterPro: IPR001264; Transglycosyl.		
DR	InterPro: IPR001460; Transpeptidase.		
DR	Pfam: PF00912; Transglycosyl. 1.		
DR	Pfam: PF00905; Transpeptidase; 1.		
DR	ProDom: PD001895; Transglycosyl. 1.		
FT	NON_TER	719	719
SO	SEQUENCE	719 AA; 79710 MW; D90EB42D5380D9E1 CRC64;	

Query Match	93.1%	Score 3245;	DB 2;	Length 719;
Best Local Similarity	92.0%;	Pred. No. 2.1e-168;		
Matches 613; Conservative	35;	Mismatches 18;	Indels 0;	Gaps 0;

Oy	1	KIKDKKNOIADLISERRVNAQANDPPDLKAVLSIEDHREFRHCIDPTRLIGAFLN	60
Db	54	KIYDKRNOIADLISERRVNAQANDIPDLKAVLSIEDHREFRHCIDSRILIGAFLN	113
Oy	61	LOSNLSOGSTLTOOLIKLTYESTSTSDOTISRAQOAMLAIOLEOKATKOELLTYINK	120
Db	114	LOSNLSOGSTLTOOLIKLTYESTSTSDOTISRAQOAMLAIOLEOKATKOELLTYINK	173
Oy	121	YVMSGNGYMGOTAAQONTYQKDLNL.SLPOLALLAGMPARNOQDPYSHPEAODRNLVL	180
Db	174	YVMSGNGYMGOTAAQONTYQKDLNL.SLPOLALLAGMPARNOQDPYSHPEAODRNLVL	233
Oy	181	SEMKNOGYSIAEOYEKAVNPRTIDGLOSLKSASNTPAYMONTYKEVINQOYEETGYMLLT	240
Db	234	SEMKNOGYSIAEOYEKAVNPRTIDGLOSLKSASNTPAYMONTYKEVINQOYEETGYMLLT	293
Oy	241	TGMDVYTTVWDEAKHLMIDYNDEBYAAYPPDELOVASTIYDVSNGKVIQOLARHSSN	300
Db	294	TGMDVYTTVWDEAKHLMIDYNSQYVSPDDELQVASTYDVDSNGKVIQOLARHASN	353
Oy	301	VSFQINQAVENRBMGSTMRTITPYAPALEGYVDYATVYVHEPPYVPCGTNPVYVNMWR	360
Db	354	VSFQINQAVENRBMGSMKRTIDYAPALEGYVDYATVYVNMWRPCTSTPVPVNMWR	413
Oy	361	GYFGNITLQVLOQSRUNPAVETLNKGLNKAFTLNGLOIDVPSIHYSAISNTTESD	420

```

|||||
Db 414 AAFGNITLQYALQOOSRNVAVETLNKVGIDRAKTFNLNGIGIDYPSMHYANAISSNTTESN 473
QY 421 KYGASSEKMAAAYAFANGGYTKPMYIHKVFSQSGSEKFSNVGTRAKKETTAYMMND 480
Db 474 KÖYGASSEKMAAAYAFANGGYTKPMYIHKVFSQSGSEKFSNVGTRAKKETTAYMMNE 533
QY 481 MKKTVLTGTRNAVYLAWLPOAGKTGTSNYTDEIEINHITKSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTVLTGTRNAVYLAWLPOAGKTGTSNYTDEIEIKHITKSTGYAPDEMFVGYTRKYSM 593
QY 541 AAWTGYSNRLTPLVGNGLTVAAKVYRSMYTYLSEGSNPEDWNIPEGLYRNGEYVFNKGAR 600
Db 594 AAWTGYSNRLTPLVGNGLTVAAKVYRSMYTYLSEDTHPEDWMPDGLFRNGEYVFNKGAR 653
QY 601 STWNSAPQOOPSTESSSSSDSTSSSTPSTNNSTTTNNNTQOOSNTTTPDOONON 660
Db 654 STWNSAPQOOPSTESSSSSDSTSSSTPSTNNSTTTNNNTQOOSNTTTPDOONON 713
QY 661 POPAOP 666
Db 714 POPAOP 719

RESULT 12
Q54946 PRELIMINARY: PRT: 719 AA.
AC 054946:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA OR PBPLA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN 11
RN SEQUENCE FROM N.A.
RC MEDLINE=93010977; PubMed=1396576;
RA Martin C., Stibold C., Hakenbeck R.;
RT "Resistance of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain.";
RL EMBO J. 11:3831-3836(1992).
RN 12
RN SEQUENCE OF 310-619 FROM N.A.
RC STRAIN=8303, AND 35193;
RA Smith A.M., Klugman K.P.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN 13
RN SEQUENCE OF 320-717 FROM N.A.
RC STRAIN=117/246, #20/B98, AND #27/SHA3;
RA MEDLINE=98409715; PubMed=9736547;
RX Asahi Y., Ubukata K.;
RT "Association of a thr-371 substitution in a conserved amino acid motif
RT of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
RN 14
RN SEQUENCE FROM N.A.
RC STRAIN=SP-665;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 0:0-0(1999).
DR EMBL: X67868; CAA48068.1; -
DR EMBL: AF046230; AAC24695.1; -
DR EMBL: AB006876; BAA32072.1; -
DR EMBL: AF139883; AAD43066.1; -
DR InterPro: IPR001264; Transglycosyl.

```

```

DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR Prodom: PD001895; Transglycosyl. 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA: 79736 MM; F090F9152C834DC CRC64;

Query Match 93.1%; Score 3243; DB 2; Length 719;
Best local similarity 91.9%; Pred. No. 2,7e-168;
Matches 612; Conservative 32; Mismatches 22; Indels 0; Gaps 0;

QY 1 KIYDNKNQOLIADIGSSRRVNAQANDIPTDLVKAIVSIEDHFRFDHGRGIDTIRILGAFLRN 60
Db 54 KIYDNKNQOLIADIGSSRRVNAQANDIPTDLVKAIVSIEDHFRFDHGRGIDIRILGAFLRN 113
QY 61 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKKAQEWALAIQLEQKATKQELITVYINK 120
Db 114 LOSNSLQGGSTLTQOLIKLTFSTSDQTSKKAQEWALAIQLEQKATKQELITVYINK 173
QY 121 VYMSNGNYGMQTAQNYVYGGKDLNNLSLPOLALLAGHPQAPNOVDYPSHPAADDNRNLVY 180
Db 174 VYMSNGNYGMQTAQNYVYGGKDLNNLSLPOLALLAGHPQAPNOVDYPSHPAADDNRNLVY 233
QY 181 SEMKNOGYISAEQYERKAVNTPITDGLQSLKSASNYPAYMDNYLKEYI NOVEEETGYNLTL 240
Db 234 SEMKNOGYISAEQYERKAVNTPITDGLQSLKSASNYPAYMDNYLKEYI NOVEEETGYNLTL 293
QY 241 TGMADVNTNDQEAQKHLMDIYNTDEVYANPDELOVASTIVDSNGKVIAQLGARHOSSN 300
Db 294 TGMADVNTNDQEAQKHLMDIYNSDQYVSPDDLOVASTIVDSNGKVIAQLGARHOASN 353
QY 301 VSRFGINQAVETNRDQSGTKPITDVA PALEYGYDDSTATIVHDEPYNPCTNPVYNMNR 360
Db 354 VSRFGINQAVETNRDQSGTKPITDVA PALEYGYDDSTATIVHDEPYNPCTNPVYNMNR 413
QY 361 GYFGNITLQYALQOOSRNVAVETLNKVGIDRAKTFNLNGIGIDYPSIHSNAISSNTTESD 420
Db 414 AAFGNITLQYALQOOSRNVAVETLNKVGIDRAKTFNLNGIGIDYPSMHYANAISSNTTESN 473
QY 421 KYGASSEKMAAAYAFANGGYTKPMYIHKVFSQSGSEKFSNVGTRAKKETTAYMMND 480
Db 474 KÖYGASSEKMAAAYAFANGGYTKPMYIHKVFSQSGSEKFSNVGTRAKKETTAYMMND 533
QY 481 MKKTVLTGTRNAVYLAWLPOAGKTGTSNYTDEIEINHITKSQFVAPDELFAGYTRKYSM 540
Db 541 AAWTGYSNRLTPLVGNGLTVAAKVYRSMYTYLSEGSNPEDWNIPEGLYRNGEYVFNKGAR 600
QY 594 AAWTGYSNRLTPLVGNGLTVAAKVYRSMYTYLSEDTHPEDWMPDGLFRNGEYVFNKGAR 653
Db 601 STWNSAPQOOPSTESSSSSDSTSSSTPSTNNSTTTNNNTQOOSNTTTPDOONON 660
QY 654 PITTEPSTQOOSTYAESSSSDSTSSSTPSTNNSTTTNNNTQOOSNTTTPDOONON 713
Db 661 POPAOP 666
Db 714 POPAOP 719

RESULT 13
Q54948 PRELIMINARY: PRT: 719 AA.
AC 054948:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.

```


Db	594	AVMTGYSNRLEPIVDGFLVAAKYVRSMITTYLSEDTHPEDWTMPDGLFRNGSEFVRNGAR	653
Qy	601	STWNSPAQOQPPSTSSSSSSSDSTSSQSSSTTPSTNNSTTTNPNNNTOOSNTTPOQOON	660
Db	654	PTWPPSTTQOOSSTASSSSSSSDSTSSQSSSTTPSTNNSTTTNPNNNTOOSNTTPOQOON	713
Oy	661	POPAQ 666	
Db	714	POPAQ 719	
RESULT	15		
ID	054950	PRELIMINARY:	PRT: 719 AA.
AC	054950:		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DE	01-JUN-2001	(TREMBLrel. 17, Last annotation update)	
DT		PENICILLIN-BINDING PROTEIN 1A (FRAGMENT),	
GN		PONA.	
OS		Streptococcus pneumoniae.	
OC		Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;	
OC		Streptococcus.	
OX		NCBI_TaxID=1313;	
RN		[1]	
RP		SEQUENCE FROM N.A.	
RC		STRAIN=2039;	
RC		MEDLINE=93010977; PubMed=1396576;	
RA		Martin C., Sibold C., Hakenbeck R.;	
RT		"Relatedness of penicillin-binding protein 1a genes from different	
RT		clones of penicillin-resistant Streptococcus pneumoniae isolated in	
RT		South Africa and Spain."	
RL		EMBO J. 11:3831-3836(1992).	
DR		EMBL, X67870; CAA48070.1; -;	
DR		InterPro: IPR001264; Transglycosyl.	
DR		InterPro: IPR001460; Transpeptidase.	
DR		Pfam: PF00912; Transglycosyl; 1.	
DR		Pfam: PF00905; Transpeptidase; 1.	
DR		ProDom: PD001895; Transglycosyl; 1.	
DR		NON_TER 719	
FO	SEQUENCE	719 AA; 79874 MW; FACCFAAB30D865D4 CRC64;	

Query Match	91.6%	Score 3192	DB 2	Length 719
Best Local Similarity	90.4%	Pred. No. 1.6e165		
Matches 602	Conservative 41	Mismatches 23	Indels 0	Gaps 0
Qy	1	KIYDKRNLADI	DGSE	RVNAQANDIPDYLVAIYSIEDHREDFHGRGIDITIRILGAFLRN 60
Db	54	KIFDKNELIADL	GSERRVNAQANIPPDYVAIYSIEDHREDFHRCVDITIRILGAFLRN 113	
Qy	61	LOSNSLGGSG	STLTQOLIKLTYFSTSTSPQTSRKAQEA	MLAQLQBOKATKQELIYYINK 120
Db	114	LONNSLGGSG	STLTQOLIKLTYFSTSTSPQTSRKAQEA	MLAQLQBOKATKQELIYYINK 173
Qy	121	YVMSGNGTAA	QNYGKDLNNLSLQALLAAMPQAPN	YDYSHPHEAODRRNLVL 180
Db	174	YVMSGNGTAA	QNYGKDLRELSTLPOLLALAMPQAPN	YDYSHPHEALDRRNLVL 233
Qy	181	SEMKNOGYS	AEUYKAVNPTDGLQSLKSA	SNPAYMDNLYKEVINOVEETGYNLLT 240
Db	234	SEMKNOGYS	AEUYKAVNPTDGLQSLKSA	SNPAYMDNLYKEVINOVEETGYNLLT 293
Qy	241	TGMDYVTVND	EAPQKHLMDIYNTDEYVAYVPDELO	VASTIYDVNSGKVAIQAQGARHOSN 300
Db	294	TGMDYVTVND	EAPQKRLMDIYNSDDYVYTPDDLO	VASTYDVNSGKVAIQAQGARHOSN 353
Qy	301	VSFGINQAV	ETNRDMSGTMKPTTDYAPALE	GYVDYSTATIYHDEBYNYPGTNPYMYNR 360
Db	354	VSFGINQAV	ETNRDMSGTMKPTTDYAPALE	GYVDYSTATIYHDEBYNYPGTNPYMYNR 413
Qy	361	GYFGNITLQ	YALQOSRNPAYETLNK	VGLNRAKTYFLNGLIDYPSIHYSNATISNTTESD 420

Db	414	VYFGNITITQVALQOQSRNVTAVETLINTKVGLDBAKFTFLNGLGIDVPSMHYANNAISSNTESN	473
Qy	421	KKYGASSSEKMAAATAAFANGSTYYKKPMYIKHVVPSDQSEKFEFNVGRAKKETTAYMTD	480
Db	474	KOYGASSSEKMAAATAAFANGSIYHKPMYINKYIVSDSEKEFPDAGRKETTAYMTTE	533
Qy	481	MKKTUYLTGTGRNMYLAWLPQAGKFTGTSNYDEEIEHNIKTSQVAPADELFAGYTRKYSM	540
Db	534	MKKTUYLADGIGRGAYLPLPQAGKFTGTSNTDDEIEKYINKTGYVAPDEMFVGYTRKYSM	593
Qy	541	AWMTGYSNRKLPLVYNGNLTVAAKYRRSMYTLTSGSNPEDMNIPBGLYRNGEYFEKKGAR	600
Db	594	AWMTGYSNRKLPLVYGDGLTVAAKYRRSMYTLTSGSNPEDMNIPBGLYRNGEYFEKKGAR	653
Qy	601	STWNSPAQOQPPSTESSSSSSDSSSTQSSSTPTSTNNSTTTNPNNNTQQOQNTTPDOQNON	660
Db	654	STWNSPAQOQPPSTESSSSSSDSSSTQSSSTPTSTNNSTTTDPNNNTQQOQNTTPDOQNON	713
Qy	661	POPAPQ 666	
Db	714	POPAPQ 719	

Search completed: June 13, 2002, 08:45:21
Job time: 252 sec

THIS PAGE BLANK (USPTO)

Db 1 KIYDNKNOLLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDRHGIDTIRILGAFLEFN 60
QY 61 LOSNSLOGSSTLTLOOLIKLTFTSTSDOTSIRKAOEAMLAIOLEKATKOELLTYINK 120
Db 61 LOSNSLOGSSTLTLOOLIKLTFTSTSDOTSIRKAOEAMLAIOLEKATKOELLTYINK 120
QY 121 VYMSNGNGMOTAAQNYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHPBAADRRLVYL 180
Db 121 VYMSNGNGMOTAAQNYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHPBAADRRLVYL 180
QY 181 SEMKNOGYISAQOYERAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEEETGYNLLT 240
Db 181 SEMKNOGYISAQOYERAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEEETGYNLLT 240
QY 241 TGMDEVYTNDOEAQKHLMDIYNTDEVYVAPDELOQVASTIVDSNGKVAIOGARHOSSN 300
Db 241 TGMDEVYTNDOEAQKHLMDIYNTDEVYVAPDELOQVASTIVDSNGKVAIOGARHOSSN 300
QY 301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPYNPGTNPVYNNMDR 360
Db 301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPYNPGTNPVYNNMDR 360
QY 361 GYFGNITLQYALQOQRNVPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSAISSNTTESD 420
Db 361 GYFGNITLQYALQOQRNVPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSAISSNTTESD 420
QY 421 KRYGASSEKMAAAYAAAFANGGTYIKPMYIHKVYFSDGSEKESNVGTRAKKETTAYMMTD 480
Db 421 KRYGASSEKMAAAYAAAFANGGTYIKPMYIHKVYFSDGSEKESNVGTRAKKETTAYMMTD 480
QY 481 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSOFVAPDELFACTYRKYSM 540
Db 481 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSOFVAPDELFACTYRKYSM 540
QY 541 AWMTGYSNRLTPLVGNGLTVAAYKVRSMNTYLSSEGSNPEDMNIPEGLYRNGEEVFNKGAR 600
Db 541 AWMTGYSNRLTPLVGNGLTVAAYKVRSMNTYLSSEGSNPEDMNIPEGLYRNGEEVFNKGAR 600
QY 601 STWNSPAPQOPPSTESSSSSDSTSSSSTTPTNNSTTTNNNTTQOOSNTTPOOQONN 660
Db 601 STWNSPAPQOPPSTESSSSSDSTSSSSTTPTNNSTTTNNNTTQOOSNTTPOOQONN 660
QY 661 POPAOP 666
Db 661 POPAOP 666

RESULT 2
US-08-481-435-6
; Sequence 6, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganesb, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

Query Match 99.5%; Score 3466; DB 3; Length 682;
Best Local Similarity 99.4%; Pred No. 5 4e-254;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNOLLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDRHGIDTIRILGAFLEFN 60
Db 17 KIYDNKNOLLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDRHGIDTIRILGAFLEFN 76
QY 61 LOSNSLOGSSTLTLOOLIKLTFTSTSDOTSIRKAOEAMLAIOLEKATKOELLTYINK 120
Db 77 LOSNSLOGSSTLTLOOLIKLTFTSTSDOTSIRKAOEAMLAIOLEKATKOELLTYINK 136
QY 121 VYMSNGNGMOTAAQNYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHPBAADRRLVYL 180
Db 137 VYMSNGNGMOTAAQNYKGDNLNLSLPOLALLAGMPQAPNOYDPYSHPBAADRRLVYL 196
QY 181 SEMKNOGYISAQOYERAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEEETGYNLLT 240
Db 197 SEMKNOGYISAQOYERAVNTPITDGLQSLKSASNPAYMDNLYKEVINOVEEETGYNLLT 256
QY 241 TGMDEVYTNDOEAQKHLMDIYNTDEVYVAPDELOQVASTIVDSNGKVAIOGARHOSSN 300
Db 257 TGMDEVYTNDOEAQKHLMDIYNTDEVYVAPDELOQVASTIVDSNGKVAIOGARHOSSN 316
QY 301 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPYNPGTNPVYNNMDR 360
Db 317 VSEFGINQAVETNRDMGSTMKPITDYAPALEYGYDSTATIVHDEPYNPGTNPVYNNMDR 376
QY 361 GYFGNITLQYALQOQRNVPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSAISSNTTESD 420
Db 377 GYFGNITLQYALQOQRNVPAVETLNKVGILNRAKTFPLNGIGIDYPSIHYSAISSNTTESD 436
QY 421 KRYGASSEKMAAAYAAAFANGGTYIKPMYIHKVYFSDGSEKESNVGTRAKKETTAYMMTD 480
Db 437 KRYGASSEKMAAAYAAAFANGGTYIKPMYIHKVYFSDGSEKESNVGTRAKKETTAYMMTD 496
QY 481 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSOFVAPDELFACTYRKYSM 540
Db 497 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSOFVAPDELFACTYRKYSM 556
QY 541 AWMTGYSNRLTPLVGNGLTVAAYKVRSMNTYLSSEGSNPEDMNIPEGLYRNGEEVFNKGAR 600
Db 557 AWMTGYSNRLTPLVGNGLTVAAYKVRSMNTYLSSEGSNPEDMNIPEGLYRNGEEVFNKGAR 616
QY 601 STWNSPAPQOPPSTESSSSSDSTSSSSTTPTNNSTTTNNNTTQOOSNTTPOOQONN 660
Db 617 STWNSPAPQOPPSTESSSSSDSTSSSSTTPTNNSTTTNNNTTQOOSNTTPOOQONN 676
QY 661 POPAOP 666
Db 677 POPAOP 682

```
RESULT 3
US-08-245-511-4
; Sequence 4, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-245-511-4

Query Match          46.5%; Score 1621; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 6.7e-115;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
DB 305 VSEGINOAVETNRDNG 320
|||||
RESULT 4
US-08-600-993A-4
; Sequence 4, Application US/0860093A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 320 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-600-993A-4

Query Match          46.5%; Score 1621; DB 2; Length 320;
Best Local Similarity 99.7%; Pred. No. 6.7e-115;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Db 185 SEMKNGYISABOYEKAVNPTITDGLQSLKASNPAYMDNLYKEVINOVEERTYNLT 244
 QY 241 TGMADVYTNDOEAKHLMIDYNTDEVAYPDDDELQVASTIVDSNGKYAOLGARHOSN 300
 Db 245 TGMADVYTNDOEAKHLMIDYNTDEVAYPDDDELQVASTIVDSNGKYAOLGARHOSN 304
 QY 301 VSGFINQAVETNRDMG 316
 Db 305 VSGFINQAVETNRDMG 320

RESULT 5
 US-08-731-716-2
 ; Sequence 2, Application US/08731716
 ; Patent No. 5789202

GENERAL INFORMATION:
 APPLICANT: Hoskins, John
 APPLICANT: Jaskunas, S. Richard
 APPLICANT: Rockey, Pamela K.
 APPLICANT: Zhao, Genshi
 APPLICANT: Rostock, Paul R., Jr.
 APPLICANT: No. 5789202 is, Franklin H.
 TITLE OF INVENTION: Penicillin Binding Protein From
 TITLE OF INVENTION: Streptococcus Pneumoniae
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Eli Lilly and Company
 STREET: Lilly Corporate Center
 CITY: Indianapolis
 STATE: Indiana
 COUNTRY: U.S.
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/731,716
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Webster, Thomas D.
 REGISTRATION NUMBER: 39, 872
 REFERENCE/DOCKET NUMBER: X-10, 887
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-276-3334
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 731 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-731-716-2

Query Match 18.3%; Score 636; DB 1; Length 731;
 Best Local Similarity 27.7%; Pred. No. 8.2e-40;
 Matches 192; Conservative 108; Mismatches 252; Indels 142; Gaps 19;

QY 2 IYDNKNQIADIAGSERRNAQANDIPTDLVAKAIVSIEDHRFEDHGIDITIRILGAFRLNL 61
 Db 95 IDPREKEKAGALSGQKVELTLDISKNLONNAVIAITEDRSFYKNDGINGRF--FLAIV 151
 QY 62 QSNLSLGGSTLQOLIKLYFSTSDQITISRAQAMLAIOEQKATQOELITYINKY 121
 Db 152 TAGRSGGSTIQLQAKNLYS--ODQYVERRAKFEFFALISLRSKYSKEQILTMYLMA 208
 QY 122 YMSNGYMGQTAQANYGKDLNLNLSPOLALLAGMPQAPNOYDPSYHPPEAODRNLVLS 181
 Db 209 YFGNGYMGVEDASKTYFGVSASEVSLDQAATTLAGMLKGPFLYPLNLSVEDSTTRRDIVLQ 268
 QY 182 EMKNQGYISABOYEKAVNPTIT---DGLQSLKASNPAYMDNLYKEVI--NOVEERT 234

Db 269 NNVAAGYIDKNOETEAQVDMKTSQLDHKYEGKISDRYRYSYFPAVAVNEAVSKYNLTREE- 327
 QY 235 GYNLFTTGMADVYTNDOEAKHLMIDYNTDEVAYPDDDELQVASTIVDSNGKYAOL 292
 Db 328 ---IYVNGYRITELDONQANMOIYENTSLFPRADGTFPQSSGVALEPTGVRGVY 384
 QY 293 GARHOSNVSF-GINQAVETNRDMGSTMKPITDYAPALEYGYDYATYIVHDEPY----- 346
 Db 385 GOVADNDKGTGFNFNFNATQSKRSPGSTIKPLVYTPFAVEGALNQLDNHTWQYDSYRV 444
 QY 347 -NYPGTNT---PYVNMWDGYEGNITLQYALQOSRVPAVETLNKGLNRKAKFTLNGLGI 401
 Db 445 DNAGIKTSREVPYIO-----SLAESLNDPAVATVNDGVCKA----- 482
 QY 402 DYPSIHYNSNAISSNTTESDKKYGASSEK-----MAAAYAPANGGT 442
 Db 483 -----FEAGEKFGLNMEKXDRVLGVALGSGOVETNPLOMNAVAFAFNEGI 527
 QY 443 YKPKMYIHKVVEPSDGE-KEFSNVGTRAKETTYAMMTDMKTVLTYGGRNAYLAMLPQ 501
 Db 528 MPEAHFISRIENASGVIAHSHKNSQKRVIDKSVADKMTSMIGTFTNGTGISSSPADYVA 587
 QY 502 AGKTGTSNTDEIEHHTSQFVAPDELFAGYTRKYSMAVWTGY-----SNRLFTLVCG 556
 Db 588 AGKTGTT---EAVFNPETYS---DQVYIGTTPDVVISHWLGFPPTDENHILASTSN 638
 QY 557 GLTVAKYRSMYTYLSEGSNPEDMNIPEGLYRNGEFVRKNGARSTWNSPAPQOPSTES 616
 Db 639 G---AAHVERINANTIL-----PYTGG 657
 QY 617 SSSSSDSTSSQSSSTPTSTNNSTTTPNNTQOS 650
 Db 658 STFTVENAYKONGIAPANTKRQYQTDNDSQTDON 691

RESULT 6
 US-08-481-435-2
 ; Sequence 2, Application US/08481435
 ; Patent No. 6027906
 GENERAL INFORMATION:
 APPLICANT: Balganes, Tanjore S
 APPLICANT: Town, Christine
 TITLE OF INVENTION: No. 6027906e1 Polypeptides
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: White & Case
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: United States
 ZIP: 10036-2787
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/481,435
 FILING DATE: 10-JUL-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: IN 580/MAS/94
 FILING DATE: 01-JUL-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: SE 9404072-2
 FILING DATE: 24-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Steiner, Richard J.
 REGISTRATION NUMBER: 35,372
 REFERENCE/DOCKET NUMBER: 1103326-151
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 819-8783


```

QY 237 NLTGMDVYTVNDOEQ-----KH-----LMDIYN-----TDEYVA 268
DB 286 RIYTT---ITRRVQQAQAQVARNRNLVDYDMRGYRGPANVLKVGSAAMDNKKITDTLKA 342
QY 269 YP-----DDEL 274
DB 343 LPTGYPLPAVATSANPDQATMLADGSTVALSMEGVWRAPRYRSDPTQGPPIPRKVTPL 402
QY 275 Q-----VASTIVDS--NGKVIOLGA--RHOSNVSFGINOAV 309
DB 403 QTGQOIWROVQDAMWLAQVPEVNSALVSIPOQNGAVMALVGGFENQSK-----FNRAI 457
QY 310 ETRNRMGSTMKPITPYAPALEXGVYDSTATIVHDEPNYPGNTPYNNDRG----- 361
DB 458 QALROVGSNIKPEL-YTAMDKGL--TLASMLND-----VPISRMDASAGSDWQPK 505
QY 362 -----YFGNITLOVALOOSRNPVAVETLKNKVLNRAKTFNLGLGIDPSIHSMAISSMT 416
DB 506 NSPPQYAGPIRLROGLGOSKNVVMRAMRGVDYAAEYLORFGFPFPAONIVHTESLA--- 562
QY 417 TESDKKYGASS--EKMAAAYAFANGCTYKKPMYIHKYVFDG----- 457
DB 563 -----LGSASTPMOVAGYAVMANGFLVDPWFISKIENDQGVIFPAKPKVACPECD 616
QY 458 -----SEKEFSNV-----GTRAKETAYM-MTDMKTVLTYYGTGRNAY----- 495
DB 617 IYVIGDPOKSNVLENNDEVDVAISREOQNSVPMPOLEQANQALVAKGADEYAPHVIN 676
QY 496 --LAMLPOA-----GKTGTSNYTDEIENHKTSGF 524
DB 677 TPLALFLIKALMTNIFGEPMOGTGRAGRDILQRDICKGTGTTN----- 721
QY 525 VAPDELFGYTKRKYSMAVVTGYSNRLTPLYGNGLYTAAVYYSMMKTYLSEGSNPEDMNIP 584
DB 722 SKDMFSGYGVVTSVWIGFDHRRNL--GHTTASGAILDQISGYEGGA----- 770
QY 585 EGLYRNGEYFVKNGARSTWNS-----PAPQOP--PTSESSSSSSDSTSGSSSTTPTST 635
DB 771 -----KSAQOPAMDAYMKAVLEGVPEQPLTPPGIYTVNIDRSTGQ----- 810
QY 636 NNSTTTNNNTQQ---SNTTPDQ 657
DB 811 ----LANGNSREYFIECTOPTQQ 831

```

RESULT 8

US-08-481-435-4

; Sequence 4, Application US/08481435

; Patent No. 6027906

; GENERAL INFORMATION:

; APPLICANT: Balganes, Tanjore S

; TITLE OF INVENTION: No. 6027906el Polypeptides

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESS: White & Case

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: United States

; ZIP: 10036-2787

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/481,435

; FILING DATE: 10-JUL-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: IN 580/MAS/94

```

; FILING DATE: 01-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Stener, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 823 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-481-435-4

```

Query Match

15.1%; Score 526; DB 3; Length 823;

Best Local Similarity 28.9%; Pred No. 2.1e-31;

Matches 203; Conservative 99; Mismatches 273; Indels 128; Gaps 30;

```

QY 9 LIALDGSERRVNAQANDIFPTDLVKAIVSIEDHREFPDHGRIDITRIGAFRLNLS-NSIQ 67
DB 183 MISSPGEQRLFPVPRSGFDLLVDLTATEDRHFEYHDSISYIGRAVLANTAGRTVQ 242
QY 68 GGSITLTOOLIKLTYSTSDTISRKAQOAMLAIOLEKATKQELTYINKYK---- 123
DB 243 GASTLQOOLVKNLFUS---SERSYWRKANEAAYMALIMDARYSDRLLEYMEVYLGQSG 299
QY 124 SNGNMGQTAQNYGKDLNLSLPOLALLAGPOAPNQYDPSHPEAADRNLVLSM 183
DB 300 DNEIRGPIPLASLYFRPVEELSLDQALLVGVKASITNWRNKLALERRNLVRLI 359
QY 184 KNOGYISABQYKAVNPTITDGLSKSA-SNPAYMDVYLKVI---NOVEETGYNL 238
DB 360 OOOQIIDLQELDYMLSARPL--GVQPRGVYISPPAPMQLVROELQAKLDKVDLSGVKI 417
QY 239 LTTGMDVYTVNDOEA-----QKHLMDIYNDEYVAYDDDLQVASTIVDS 284
DB 418 FTT---FDSVAQDAEKAKEAVGIPALKKORKISD-----LETAIVVDHF 459
QY 285 NGKVIOLGARHOSNVSE-GINQAVETNRDMGSTMKPTDYAPALEYGVYDSTATIVHD 343
DB 460 SGEVRAMVG-----GSEPPQAGYNRRAMQARRSIGSLAKPAT-YLTALSQKIYRLNWTWD 514
QY 344 EPNYVPGTNPYV--NMDRGY--FGNITLOVALOOSRNPVAVETLKNKVLNRAKTFLLNG 398
DB 515 APIALRQPNQGVWSPQNDNRYSSEGRVMLVDALTRSMNVPTVNLGMLGLPAVETWIK 574
QY 399 LGIDYPSIHSNAI---SNTTESDKKYGASSEKMAAATAARANGTYKKPIHKYVS 455
DB 575 LGVPKQDLHPVAMLIGALNLTPIE-----VAQAQOTIASGGRAPLSALRSVIAE 625
QY 456 DGS--EKEFSNVGTRAMKETFTAYMTDMKTYLTVGTGR--NAVYLAWLQAGTGSNT 511
DB 626 DQKVLQSFQ--AERAVPQAAVILITMTQOYVQKSTGOLGAKKYINLILAGTGTNN-- 682
QY 512 DEIENHIKTSQVAPDELFGYTRKYSMAVVTGYSNRLTPLYGNGLYTAAKYSRSMY 571
DB 683 -----NNV-----DTMEFAGIDGSTVITVWGRDNN-QPTKLVGASGAMSIYQ---RY 725
QY 572 LSEGSNPEDMNI--PE-----GLYRNGEYFVKNGAR--STWNSPAPQOPPTSESSSSSD 622
DB 726 LA-NQTPPLNLVPPEDIADMGVDYDGNFVCSGGMILPVMVS-----DPSLCOQ 775
QY 623 SNTSSSTTPTNNSTTTNPNNTQQSNTTTPDOONONPOPA 665
DB 776 SEMQO-----QPSGNPFDQSSQPOQDPOQ-QPAQ 803

```


RESULT 9
US-08-481-435-7
Sequence 7, Application US/08481435
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balganes, Tanjore S
APPLICANT: Town, Christine
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481.435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 844 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
IMMEDIATE SOURCE:
CLONE: PARC0438 PBP 1B Q0AA
US-08-481-435-7

Query Match 14.8%; Score 514; DB 3; Length 844;
Best Local Similarity 28.6%; Pred. No. 1.8e-30;
Matches 201; Conservative 99; Mismatches 275; Indels 128; Gaps 30;

QY 9 LIADGSRNRNAQNDPTDLVKAIVSIEDHREPDHGRGIDPTIRILGAFRLNLSQ 67
DB 204 MISSPNGRQLFVRSGPDLVLTLLTEDRHVEHGDISISGRVLANLRAGRVQ 263
QY 68 GSGTLTQOLILKYFSTSDQTSIRKQAQEMLAQLQEKATKQELITYYINKYVM--- 123
DB 264 GASTLTAALVNLFLS---SERSYRKANEAVMALIMDARYSKDILLEYNNVEYLGSG 320
QY 124 SNGNTGMOATNYGKLNLSLPQALLAGMPQAPQYDPYSHEPAQDRNLVLSEM 183
DB 321 DNEIRGFPLASLYYFGRVEEELSDQALLVGMVKGASITYNWRPKLALERRNLVLRLL 380
QY 184 KNOGYISAQYKAVNPTITGLOSLKSA-SNYPAYMNYLKEVY---NOVEETGYNL 238
DB 381 QQQQLITDELIDMLSARPL--GVQPRGVYISQPAFMOLVRELOAKGDKKIDSGVKT 438
QY 239 LTTGMDVYTNVQEA-----OKHLMIDYNTDEVAVYDDELQVASTIVVS 284

DB 439 FTT-----FDSVADQAEKAAVEGIPALKKORRLSD-----LETAIVVDRE 480
QY 285 NGXYIAQLGARRQSSNSGF-GINQAVENNRDMGSMKPTTQAPALAEVYSTFTIYHD 343
DB 481 SGEVRAWVG---GSEPOFAGYNRAMQARRSIGSLAKPAT-YTLALSQPKIYRLMTIAD 535
QY 344 ERYNPGTNTPY---NMDRGY--FGNITLQYALQOSRVPVAVETLUNKGLNRAKTFUNG 398
DB 536 APIALRQNGQVWSPQNDKRSSEGRVNLVDALTRSNANVPYVNLGMALGLPAVETWK 595
QY 399 LGIDPYSIHYNSAI---SSNTESDKRYGASSEKMAAAYAAPANGYTKYKPYIKHYVFS 455
DB 596 LGVPKQDLHPVAPMLGLALNLTPIE-----VAQAFOTIASGNRAPLSALRSVIAE 646
QY 456 DGS--EKEESNNGTRAMKETTAYMMTDMMKYVLYTGTR--NAYLAMLPOAGKTSTNNT 511
DB 647 DGVLYQSFPQ-AERAVPAQAAVLTLMWQOYVQKGTGLQAGAKYPNLHLAKTKTITN-- 703
QY 512 DEIEIHITKTSQFVAPDELFAGYTRKYSMAVVTGYSNRLTPLVNGQLVAAKVRSMYTY 571
DB 704 ---NNV-----DWFACIDGSTVYITWVGKDN--QTKIKLGASGAMSIYQ---RY 746
QY 572 LSEGSNPEDWNT--PE-----GLYRNGEFVFNKAR--STWNSPAPQOPPSTESSSSSD 622
DB 747 LA-NQTPPLNLVPPEDIAIDMGVDYDGNFVCSGGRILIPVMTS-----DPSLCOQ 796
QY 623 STSOSSTTPTSTNTTTPNNNTQOQSTTPDQONQPNQ 665
DB 797 SEMQ-----QPSGNPDDSSQPOQPOQ-QPQ 824

RESULT 10
US-08-481-435-8
Sequence 8, Application US/08481435
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balganes, Tanjore S
APPLICANT: Town, Christine
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481.435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9404072-2
FILING DATE: 24-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 844 amino acids

; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; ORGANISM: Escherichia coli
 ; IMMEDIATE SOURCE:
 ; CLONE: PARC0468 pBP 1B QOOL
 ; US-08-481-435-8

Query Match 14.7%; Score 512; DB 3; Length 844;
 Best Local Similarity 28.6%; Pred. No. 2.5e-30;
 Matches 201; Conservative 99; Mismatches 275; Indels 128; Gaps 30;

```

QY 9 L IADLSERRVNAQANDIPTDLVKAIVSTEDHREFDHRCIDTIRILGAFRLNLSQ--NSLQ 67
DB 204 MISSPGEORLFLVPRSGFPDLVDLTLLATEDRHFYEDHDSISYSGRAVLANTLAGRTVQ 263
QY 68 GGSTLTQOLIKLTYFSTSDQTSRKAQEAFLAIOLEOKATKOEILTYINKVYM----123
DB 264 GASTLFLLLVKNLFLS---SERSYWRKANEAVALIMDARYSKDRILELMNEVYLGQSG 320
QY 124 SNGNYGMQTAQNYTKGDLNLSLPOLALLAGPQAPNOYDPSHPBAADRNLVISEM 183
DB 321 DNEIRGFPLASLYFGRPVEELSDQOALLGVKASISYNPRNPKLALERENLRLRL 380
QY 184 KNGYISAEQYKAVNTPITDGLQSLKSA--SNYPAYMDNTLKEVI---NOVEETGYNL 238
DB 381 QOOOIIDQELYDMLSARPL--GVQPRGVIYSPQAFMQLVKROLOAKLGDKVADLSGVKI 438
QY 239 LTTGMDVYTNVDOEA-----OKHLMIDIYNTDEVYVAPDDELQVASTIVDVS 284
DB 439 FTT----FDSVADDAEKAKEAVEGIPALKKORKLSD-----LETAIVVVDRE 480
QY 285 NKVITAOGLARHSSNVSE--GINQAVETNRDMSGTMRKPTITDAPALEGYVDSATIVHD 343
DB 481 SGSEVRAMVG-----GSEPPQAGYNRAQARRSIGSLAKPAT--YLTALSOPIRYRLNTWID 535
QY 344 EPPYNYPGTNTPVY---NMDRGY--FGNITLYQVLAQSRNVPAYETLNKYGILNRAKTFPLNG 398
DB 536 APIALMQPRGQWVSPQNDRRYSSEGRVMLYDALTRSMVPTVYNLGMLGCLPAVTEWIK 595
QY 399 LGIDVPSIHYSNAI---SSNTSESDDKKYGAASSEKMAAAYAANGSTYYKKPMYIHKVPS 455
DB 596 LGVPRKQDLHPVAMLLGALNLTPIE-----VAQAFQTIASGGRNAPLSALRSVIE 646
QY 456 DGS--EKESNNGTMRAMKETIAYAMTDMKTYLFTYGTGR--NAYLAMLFQAGCTGTSINT 511
DB 647 DGVLYTOSFPQ--AERAVPAQAALTLTMTQOQVVOQRTGROLGAKYPNLHLAAGTGTN-- 703
QY 512 DEIEHNITSOFAVDELFACTRYKYSMAVNTGYSNRLTPIVANGNLJVAAKYRSMMTY 571
DB 704 -----NNV-----DTMFAGIDGSTVITVWGRDNN--QPRKLYIASGASMSIYQ--RY 746
QY 572 LSEGSPEDEMTN--PE-----GLYRNGEFVFNKNGAR--STWNSPAPQPPSTESSSSSD 622
DB 747 LA--NQTPTPLNLVPPEDIADMGVDYDGNFVCSGMRILPVMTS-----DPQSLCOQ 796
QY 623 SSTSSSTPTSTNNSTTTPNNNTQOQSTPTDQONQONQOPAQ 665
DB 797 SEMQO-----QPSGNPPOSSQPOQOPOQ--QPAQ 824

```

RESULT 11

; US-08-481-435-9
 ; Sequence 9, Application US/08481435
 ; Patent No. 6027906
 ; GENERAL INFORMATION:
 ; APPLICANT: Balganes, Tanjore S
 ; APPLICANT: Town, Christine
 ; TITLE OF INVENTION: No. 6027906el Polypeptides
 ; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: White & Case
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States
 ; ZIP: 10036-2787
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/481,435
 ; FILING DATE: 10-JUL-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: IN 580/MAS/94
 ; FILING DATE: 01-JUL-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: SE 9404072-2
 ; FILING DATE: 24-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Steiner, Richard J.
 ; REGISTRATION NUMBER: 35,372
 ; REFERENCE/DOCKET NUMBER: 1103326-151
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 819-8783
 ; TELEFAX: (212) 354-8113
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 836 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; ORGANISM: Escherichia coli
 ; IMMEDIATE SOURCE:
 ; CLONE: PARC0469 pBP 1B del 8
 ; US-08-481-435-9

Query Match 14.0%; Score 489; DB 3; Length 836;
 Best Local Similarity 28.1%; Pred. No. 1.3e-28;
 Matches 197; Conservative 97; Mismatches 274; Indels 134; Gaps 30;

```

QY 9 L IADLSERRVNAQANDIPTDLVKAIVSTEDHREFDHRCIDTIRILGAFRLNLSQLOG 68
DB 204 MISSPGEORLFLVPRSGFPDLVDLTLLATEDRHFYEDHDSISYSGRAVLANTLAGR--- 260
QY 69 GSTLTQOLIKLTYFSTSDQTSRKAQEAFLAIOLEOKATKOEILTYINKVYM----S 124
DB 261 ----TYQVLKNLFLS---SERSYWRKANEAVALIMDARYSKDRILELMNEVYLGQSGD 313
QY 125 NNGYGMQTAQNYTKGDLNLSLPOLALLAGPQAPNOYDPSHPBAADRNLVISEMK 184
DB 314 NEIRGFPLASLYFGRPVEELSDQOALLGVKASISYNPRNPKLALERENLRLRLQ 373
QY 185 NCGYISAEQYKAVNTPITDGLQSLKSA--SNYPAYMDNTLKEVI---NOVEETGYNL 239
DB 374 QOOOIIDQELYDMLSARPL--GVQPRGVIYSPQAFMQLVKROLOAKLGDKVADLSGVKIF 431
QY 240 TTGMDVYTNVDOEA-----OKHLMIDIYNTDEVYVAPDDELQVASTIVDVS 285
DB 432 TT----FDSVADDAEKAKEAVEGIPALKKORKLSD-----LETAIVVVDRE 473
QY 286 GVIYAQGLARHSSNVSE--GINQAVETNRDMSGTMRKPTITDAPALEGYVDSATIVHD 344
DB 474 GEYRAMVG-----GSEPPQAGYNRAQARRSIGSLAKPAT--YLTALSOPIRYRLNTWID 528
QY 345 PYNYPGTNTPVY---NMDRGY--FGNITLYQVLAQSRNVPAYETLNKYGILNRAKTFPLNG 399

```

```

Db 529 PIALQPNQGVSPQNDRRRSESGRWLVDAITRSNVPVINGMALGPAVETWIKL 588
Qy 400 GIDYPSIHYNAI--SSNTTESDKKYGASSEKMAAAYAFANGSTYRKPMYIHKVPSD 456
Db 589 GVPKQDLIPVAMLIQALNLTPIE-----VAQAFOTIASGGRNPLSLRLRVIED 639
Qy 457 GS--EKESNNGTRAMKETTYAMTDMKKIVLYTGTGR--NAYLAWLPQAGTGTSTNTD 512
Db 640 GKVLQSPFO--AERAVPQAQAYLLTMTQOVVORCTGRQLGAKYFNHLHLAGTGTN--- 695
Qy 513 EEIENHITSGFVAPDELFCAGTRKRYSAVMYTGYSNRLTPLVGNGLTVAAKVYRSMTYL 572
Db 656 -----NNV-----DIWFAGIDGSTVTITWVGRDN--QPTKLYGASGAMSITQ---RYL 739
Qy 573 SEGSPEDMNI--PE-----GLYNGEVEFKNRGAR--STWNSPAPQOPSTESSSSSDS 623
Db 740 A-NQTPPLNLYVPEDIDMGVDYDGNVFCGGMRLPLVMTS-----DPSLCOQS 789
Qy 624 STSQSSSTPSTNNSTTNPNNNTQOQNTPTPDQONQNPQPAQ 665
Db 790 EMQ-----QPSGNPFDOSSQPPQOPQO-QPAQ 816

```

```

RESULT 12
US-08-245-511-24
; Sequence 24, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6

```

```

; IMMEDIATE SOURCE:
; CLONE: SPRU42
; US-08-245-511-24

```

```

Query Match 11.7%; Score 406; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 6,2e-24;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 240 TTGMDVYTNVDOEAKHLMIDYNTDEYVAYPPDELOVASTIYDVSNGKVIQOLGARHOSS 299
Db 1 TTGMDVYTNVDOEAKHLMIDYNTDEYVAYPPDELOVASTIYDVSNGKVIQOLGARHOSS 60
Qy 300 NVSEGINQAVETNRDNG 316
Db 61 NVSEGINQAVETNRDNG 77

```

```

RESULT 13
US-08-600-993A-24
; Sequence 24, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 77 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; STRAIN: R6
; IMMEDIATE SOURCE:

```

; CLONE: SPRU42
US-08-600-993A-24

Query Match 11.7%; Score 406; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 6.2e-24;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 240 TTGMMVYTVNDQAEQKHLMDIYNTEDEYVAYPPDELOVASTIVDSNGKYIAQLGARHSS 299
Db 1 TTGMDVYTVNDQAEQKHLMDIYNTEDEYVAYPPDELOVASTIVDSNGKYIAQLGARHSS 60

Oy 300 NVSEGINAVETNRDMG 316
Db 61 NVSEGINAVETNRDMG 77

RESULT 14
US-08-481-435-12
; Sequence 12, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Steiner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; IMMEDIATE SOURCE:
; CLONE: PARC 0593 truncated soluble PBP 1B
US-08-481-435-12

Query Match 10.6%; Score 370.5; DB 3; Length 532;
Best Local Similarity 30.7%; Pred. No. 6.2e-20;
Matches 115; Conservative 68; Mismatches 138; Indels 53; Gaps 12;

Qy 9 LIADIGSERVNAQAANDIPTLVKAIVSIEDHREFDHRGIDITRILGAFLRNLOS-NSLQ 67
Db 183 MISSPNEGQRLFPVRSRGFDLLVDLTATEDRHFYEHDCISLYSIGRAVIANLTAGRTVQ 242

Qy 68 GGSSTLQQILKLTFFSTSDQTSKRQAEMLAIOLEKATKQELTYIINKVYM----- 123
Db 243 GASTLTLQQLVKKILFLS-----SERSYWRKANEAYMALMDARYSKDRLLLEYMNEVYLGSG 299

Qy 124 SNGNYGMOTAAQNYKGDINNLSLPOLALLAGPQAPNODPYSHPEAQDNRNLVLSM 183
Db 300 DNEIRGFPLASLYFERPVEELSLDQALLVGMKASLYNWRNRKLLERNLVLRLL 359

Qy 184 KNOGYISAEQYEKAVNPTITDGLQSLKSA-SNPAYMDNYLKEVI---NOVEETGYNL 238
Db 360 OOOQILIDQLYDMLSARPL-GVQPRGVISSPQAFMQLVROELQAKLGDKVXLSGVKXI 417

Qy 239 LTGMDVYTVNDQAE-----QKHLMDIYNTEDEYVAYPPDELOVASTIVDS 284
Db 418 FTT---FDSVAQDAEKAKEAVEGIPALKRKRLSD-----LETAIVVDRH 459

Qy 285 NGKYIAQLGARHSSNVSE-GINAVETNRDMGSTKKPTDYAPALEYGVDSSTATIVHD 343
Db 460 SEVRRAMVG-----GSEPPQAGYNRAMQARSTIGSLAKPAT-ITLALSPKTYIRLNTWIND 514

Qy 344 EPNYPPGINTPYVN 357
Db 515 APVALRQPNQVWS 528

RESULT 15
US-08-481-435-11
; Sequence 11, Application US/08481435
; Patent No. 6027906
; GENERAL INFORMATION:
; APPLICANT: Balganes, Tanjore S
; APPLICANT: Town, Christine
; TITLE OF INVENTION: No. 6027906el Polypeptides
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,435
; FILING DATE: 10-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IN 580/MAS/94
; FILING DATE: 01-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9404072-2
; FILING DATE: 24-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Steiner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-151
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

```

: MOLECULE TYPE: peptide
:
: ORIGINAL SOURCE:
: ORGANISM: Escherichia coli
:
: IMMEDIATE SOURCE:
: CLONE: PARC 0592 truncated PBP 1B
:
: US-08-481-435-11

```

Query Match	10.6%;	Score 370.5;	DB 3;	Length 553;
Best Local Similarity	30.7%;	Pred. No. 6.5e-20;		
Matches 115; Conservative	68;	Mismatches 138;	Indels 53;	Gaps 12;

[illegible]

Search completed: June 13, 2002, 08:41:58
Job time: 84 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:40:34 ; Search time 115.35 Seconds
(Without alignments)
2032.241 Million cell updates/sec

Title: US-08-961-083-2
Perfect score: 3484
Sequence: 1 KIYDNKNOIADLGSRVY.....TQSSNTTPPOQNONPOPAQ 666

Scoring table:
BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues
Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1:	/cgn2_6/ptodata/2/paa/PCNUS_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US08_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US099_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US100_COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US101_COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3484	100.0	666	19	US-09-536-784-2
2	3484	100.0	666	21	US-09-765-271-2
3	3484	100.0	666	21	US-09-765-272-2
4	3483	100.0	719	19	US-09-583-110-3973
5	3483	100.0	721	15	US-09-107-433-3705
6	3466	99.5	719	1	PCT-US02-03987-13423
7	3466	99.5	719	22	US-09-815-242-13423

8	3466	99.5	719	24	US-10-072-851-13423	Sequence 13423, A
9	1862.5	53.5	823	15	US-09-107-532-5667	Sequence 5667, Ap
10	1862.5	53.5	823	15	US-09-107-532A-5667	Sequence 5667, Ap
11	1751	50.3	778	1	PCT-US02-03987-10728	Sequence 10728, A
12	1751	50.3	778	22	US-09-815-242-10728	Sequence 10728, A
13	1751	50.3	778	24	US-10-072-851-10728	Sequence 10728, A
14	1751	50.3	789	15	US-09-134-000-4939	Sequence 4939, Ap
15	1621	46.5	320	5	US-08-116-541-4	Sequence 4, Appl1
16	1473.5	42.3	771	20	US-09-634-238-229	Sequence 229, App
17	1473.5	42.3	771	20	US-09-634-238-262	Sequence 262, App
18	1361	39.1	462	24	US-10-091-007-92	Sequence 92, Appl
19	817.5	23.5	727	1	PCT-US02-03987-5782	Sequence 5782, Ap
20	817.5	23.5	727	1	PCT-US02-03987-12777	Sequence 12777, A
21	817.5	23.5	727	22	US-09-815-242-5782	Sequence 5782, Ap
22	817.5	23.5	727	22	US-09-815-242-12777	Sequence 12777, A
23	817.5	23.5	727	24	US-10-072-851-5782	Sequence 5782, Ap
24	817.5	23.5	727	24	US-10-072-851-12777	Sequence 12777, A
25	815.5	23.4	730	20	US-09-611-529-4512	Sequence 4512, Ap
26	811.5	23.3	716	26	US-60-253-625-2437	Sequence 2437, Ap
27	811.5	23.3	716	26	US-60-257-931-2821	Sequence 2821, Ap
28	811.5	23.3	716	26	US-60-269-308-3840	Sequence 3840, Ap
29	809.5	23.2	746	15	US-09-134-001C-3214	Sequence 3214, Ap
30	661.5	19.0	675	20	US-09-634-238-228	Sequence 228, App
31	655	18.8	698	21	US-09-739-449-11825	Sequence 11825, A
32	655	18.8	698	22	US-09-803-110-11825	Sequence 11825, A
33	648.5	18.6	814	17	US-09-328-352-4373	Sequence 4373, Ap
34	641	18.4	834	16	US-09-252-991A-32671	Sequence 32671, A
35	636	18.3	731	1	US-09-107-433-4707	Sequence 4707, Ap
36	636	18.3	731	1	PCT-US97-19070-2	Sequence 4720, Ap
37	636	18.3	731	19	US-09-583-110-4720	Sequence 6216, Ap
38	632.5	18.2	855	17	US-09-328-352-6216	Sequence 845, App
39	625	17.9	660	1	PCT-US97-05223-845	Sequence 876, App
40	625	17.9	660	1	PCT-US97-05223-876	Sequence 876, App
41	625	17.9	660	11	US-08-761-318-845	Sequence 845, App
42	625	17.9	660	11	US-08-761-318-876	Sequence 876, App
43	625	17.9	660	12	US-08-824-132-845	Sequence 845, App
44	625	17.9	660	12	US-08-824-132-876	Sequence 876, App
45	625	17.9	660	13	US-08-993-002A-5603	Sequence 5603, Ap

ALIGNMENTS

RESULT 1
US-09-536-784-2
Sequence 2, Application US/09536784
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/536,784
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: OCT-30-1997
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB34063

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 666 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-536-784-2

Query Match 100.0%; Score 3484; DB 19; Length 666;
 Best Local Similarity 100.0%; Pred. No. 3e-289;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRYNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLELN 60
 DB 1 KIYDNKNQIADLGSERRYNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLELN 60
 QY 61 LOSNSLOGSSTLTQOLIKLTFSTSDOTISRKQAEAWLAIQLEOKATKOEILTYIYNK 120
 DB 61 LOSNSLOGSSTLTQOLIKLTFSTSDOTISRKQAEAWLAIQLEOKATKOEILTYIYNK 120
 QY 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAQRNRLVL 180
 DB 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAQRNRLVL 180
 QY 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAQRNRLVL 180
 DB 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAQRNRLVL 180
 QY 181 SEMKNGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLT 240
 DB 181 SEMKNGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLT 240
 QY 181 SEMKNGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLT 240
 DB 181 SEMKNGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLT 240
 QY 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVTIAQAGARQSSN 300
 DB 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVTIAQAGARQSSN 300
 QY 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVTIAQAGARQSSN 300
 DB 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVTIAQAGARQSSN 300
 QY 301 VSEFGINQAVETNRDMGSTMKPTIDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNNMDR 360
 DB 301 VSEFGINQAVETNRDMGSTMKPTIDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNNMDR 360
 QY 361 GYFGNITLQYALQOOSRNVAVELLNKVGILNRAKTFPLNGCIDYPSIHYSNALSSNTTESD 420
 DB 361 GYFGNITLQYALQOOSRNVAVELLNKVGILNRAKTFPLNGCIDYPSIHYSNALSSNTTESD 420
 QY 421 KYYGASSEKMAAAYAFANGGTYIKPMYIHKYVFSGSEKESFNVGTRAMKETTAAMMD 480
 DB 421 KYYGASSEKMAAAYAFANGGTYIKPMYIHKYVFSGSEKESFNVGTRAMKETTAAMMD 480
 QY 421 KYYGASSEKMAAAYAFANGGTYIKPMYIHKYVFSGSEKESFNVGTRAMKETTAAMMD 480
 DB 421 KYYGASSEKMAAAYAFANGGTYIKPMYIHKYVFSGSEKESFNVGTRAMKETTAAMMD 480
 QY 481 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFAGYTRKYSM 540
 DB 481 MKKTYLVYTGGRNAYLAWLPQAGKTGTSNYTDEEIEENHIKTSQFVAPDELFAGYTRKYSM 540
 QY 541 AWTGSGNLTPLVNGGLTVAAYKVRSMNTYISEGSPEDDMNIPBGLYNGEVEFNGAR 600
 DB 541 AWTGSGNLTPLVNGGLTVAAYKVRSMNTYISEGSPEDDMNIPBGLYNGEVEFNGAR 600
 QY 601 STWNSAPQPPSTESSSSSDSSTPSTNNSTTTNNNNTOOSNTTPOOQN 660
 DB 601 STWNSAPQPPSTESSSSSDSSTPSTNNSTTTNNNNTOOSNTTPOOQN 660
 QY 661 PQAPQ 666
 DB 661 PQAPQ 666

NUMBER OF SEQUENCES: 452
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/765,271
 FILING DATE: 22-Jan-2001
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/536,784
 FILING DATE: <unknown>
 APPLICATION NUMBER: 08/961,083
 FILING DATE: OCT-30-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Michelle S. Marks
 REGISTRATION NUMBER: 41,971
 REFERENCE/DOCKET NUMBER: PB340P3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 666 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-765-271-2

Query Match 100.0%; Score 3484; DB 21; Length 666;
 Best Local Similarity 100.0%; Pred. No. 3e-289;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRYNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLELN 60
 DB 1 KIYDNKNQIADLGSERRYNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLELN 60
 QY 61 LOSNSLOGSSTLTQOLIKLTFSTSDOTISRKQAEAWLAIQLEOKATKOEILTYIYNK 120
 DB 61 LOSNSLOGSSTLTQOLIKLTFSTSDOTISRKQAEAWLAIQLEOKATKOEILTYIYNK 120
 QY 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAQRNRLVL 180
 DB 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAQRNRLVL 180
 QY 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAQRNRLVL 180
 DB 121 VYMSNGNTGMQTAQNYRYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPDEAQRNRLVL 180
 QY 181 SEMKNGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLT 240
 DB 181 SEMKNGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLT 240
 QY 181 SEMKNGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLT 240
 DB 181 SEMKNGYISAEQYKAVNPTITDGLQSLKASNPAYMDNLYKEYINQVEETGYNLT 240
 QY 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVTIAQAGARQSSN 300
 DB 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVTIAQAGARQSSN 300
 QY 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVTIAQAGARQSSN 300
 DB 241 TGMADVYTNVDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDVSNKVTIAQAGARQSSN 300
 QY 301 VSEFGINQAVETNRDMGSTMKPTIDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNNMDR 360
 DB 301 VSEFGINQAVETNRDMGSTMKPTIDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNNMDR 360
 QY 361 GYFGNITLQYALQOOSRNVAVELLNKVGILNRAKTFPLNGCIDYPSIHYSNALSSNTTESD 420
 DB 361 GYFGNITLQYALQOOSRNVAVELLNKVGILNRAKTFPLNGCIDYPSIHYSNALSSNTTESD 420
 QY 421 KYYGASSEKMAAAYAFANGGTYIKPMYIHKYVFSGSEKESFNVGTRAMKETTAAMMD 480


```
|||||
Db 421 KKGASSEKMAAAVAFAANGTYYKPMYIKHVPSDSEKEFSNVGTRAKMETAYAMTD 480
QY 481 MKTVLTYGGRNAYLAMLPOAGKTGTSNTDEIEHNIKTSQVADDELFAGTRKYSM 540
Db 481 MKTVLTYGGRNAYLAMLPOAGKTGTSNTDEIEHNIKTSQVADDELFAGTRKYSM 540
QY 541 AVMTGYNRLPLVGNGLTVAAKYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
Db 541 AVMTGYNRLPLVGNGLTVAAKYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
QY 601 STWNSPAPQOPPTSSSSSSDSTSSSTPSTNNSTTNPNNNTQQSNWTPDQONON 660
Db 601 STWNSPAPQOPPTSSSSSSDSTSSSTPSTNNSTTNPNNNTQQSNWTPDQONON 660
QY 661 POPAOP 666
Db 661 POPAOP 666

RESULT 3
US-09-765-272-2
; Sequence 2, Application US/09765272
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-272-2.

Query Match 100.0%; Score 3484; DB 21; Length 666;
Best Local Similarity 100.0%; Pred. No. 36-289;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
|||||
Db 61 LQSNLSGGSTLTLQOLIKLTYFSTSTSDQTSIRKAQSAWLALQLEKATKQELITYYINK 120
QY 121 VYNSNGRYGMQTAQNTYGGKDLNLSLPOLALLAGMQAPQRYPYSHPEAQAORRLYL 180
Db 121 VYNSNGRYGMQTAQNTYGGKDLNLSLPOLALLAGMQAPQRYPYSHPEAQAORRLYL 180
QY 181 SEMNNGYISAEOYEKAVNPTITDGLQSLKSASNPAYMDYLEKEVINOVEEETGYMLLT 240
Db 181 SEMNNGYISAEOYEKAVNPTITDGLQSLKSASNPAYMDYLEKEVINOVEEETGYMLLT 240
QY 241 TGMDDVTNVDOEAQKHLMDIYNTDEYVAYPDDELQVASTIVDSNGKVIAGLARHOSN 300
Db 241 TGMDDVTNVDOEAQKHLMDIYNTDEYVAYPDDELQVASTIVDSNGKVIAGLARHOSN 300
QY 301 VSFGINQAVETNRDWSGTMKPTIDYAPALEGYDSTATTIVHDEPYNPGTNTFVYWMDR 360
Db 301 VSFGINQAVETNRDWSGTMKPTIDYAPALEGYDSTATTIVHDEPYNPGTNTFVYWMDR 360
QY 361 GYEGNITLOYALQOSRWVPVAVETLNKGLNRAKTFELNGLDIDYPSIHYSNAISSNTTESD 420
Db 361 GYEGNITLOYALQOSRWVPVAVETLNKGLNRAKTFELNGLDIDYPSIHYSNAISSNTTESD 420
QY 421 KKGASSEKMAAAVAFAANGTYYKPMYIKHVPSDSEKEFSNVGTRAKMETAYAMTD 480
Db 421 KKGASSEKMAAAVAFAANGTYYKPMYIKHVPSDSEKEFSNVGTRAKMETAYAMTD 480
QY 481 MKTVLTYGGRNAYLAMLPOAGKTGTSNTDEIEHNIKTSQVADDELFAGTRKYSM 540
Db 481 MKTVLTYGGRNAYLAMLPOAGKTGTSNTDEIEHNIKTSQVADDELFAGTRKYSM 540
QY 541 AVMTGYNRLPLVGNGLTVAAKYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
Db 541 AVMTGYNRLPLVGNGLTVAAKYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
QY 601 STWNSPAPQOPPTSSSSSSDSTSSSTPSTNNSTTNPNNNTQQSNWTPDQONON 660
Db 601 STWNSPAPQOPPTSSSSSSDSTSSSTPSTNNSTTNPNNNTQQSNWTPDQONON 660
QY 661 POPAOP 666
Db 661 POPAOP 666

RESULT 4
US-09-583-110-3973
; Sequence 3973, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3973
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3973

Query Match 100.0%; Score 3483; DB 19; Length 719;
Best Local Similarity 99.8%; Pred. No. 4,2e-289;
Matches 665; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
|||||
Db 54 KIYNKNKLIADLGSERRVNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFLRN 113
QY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDOTS SRKQOEAFLAIQLEOKATKOEILYYINK 120
Db 114 LOSNSLOGSSTLTQOLIKLTYFSTSDOTS SRKQOEAFLAIQLEOKATKOEILYYINK 173
QY 121 VYMSNGNMGMTAAQNYGKDLNLSLPQALLAGMPAPNOYDPYSHPEAAQDRRLVL 180
Db 174 VYMSNGNMGMTAAQNYGKDLNLSLPQALLAGMPAPNOYDPYSHPEAAQDRRLVL 233
QY 181 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEVINQVEEETYNLLT 240
Db 234 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEVINQVEEETYNLLT 293
QY 241 TGMDDYTNVDOBAQKHLMDIYNTDEYVAYPPDELOVASTIVDSNGKYIAQIGARHOSN 300
Db 284 TGMDDYTNVDOBAQKHLMDIYNTDEYVAYPPDELOVASTIVDSNGKYIAQIGARHOSN 353
QY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNMNR 360
Db 354 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNMNR 413
QY 361 GYFGNITLQYALQOSRNPAYETLNKVGILNRKPTFLNGLGIDYPSIHYSNATSSNTTESD 420
Db 414 GYFGNITLQYALQOSRNPAYETLNKVGILNRKPTFLNGLGIDYPSIHYSNATSSNTTESD 473
QY 421 KKYGSSSKMAAAYAFANGGTYYPKMYIHKVYFSDGSEKESNNGTRAMKETTAYMMTD 480
Db 474 KKYGSSSKMAAAYAFANGGTYYPKMYIHKVYFSDGSEKESNNGTRAMKETTAYMMTD 533
QY 481 MKKTVLTGTRNAYLAMPQAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTVLTGTRNAYLAMPQAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 593
QY 541 AVMTGYSNRLPLVNGNGLTVAAYKVRSMWTYISEGSPEDMNIPEGILRNGEFVRKNGAR 600
Db 594 AVMTGYSNRLPLVNGNGLTVAAYKVRSMWTYISEGSPEDMNIPEGILRNGEFVRKNGAR 653
QY 601 STWNSPAPQOPSTSSSSSSSDSSSTSPSTNNSTTTPNNNTTQOOSNTTPOQONN 660
Db 654 STWNSPAPQOPSTSSSSSSSDSSSTSPSTNNSTTTPNNNTTQOOSNTTPOQONN 713
QY 661 POPAP 666
Db 714 POPAP 719
```

```
RESULT 5
US-09-107-433-3705
: Sequence 3705, Application US/09107433
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stramm and David Bush
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
: SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
: THERAPEUTICS FOR DIAGN
: NUMBER OF SEQUENCES: 5206
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02354
: COMPUTER READABLE FORM:
: MEDIUM TYPE: CD-ROM ISO9660
: COMPUTER: <unknown>
: OPERATING SYSTEM: <unknown>
: SOFTWARE: <unknown>
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/107,433
: FILING DATE: 30-Jun-1998
: PRIOR APPLICATION DATA:
```

```
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3705:
SEQUENCE CHARACTERISTICS:
LENGTH: 721 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...721
SEQUENCE DESCRIPTION: SEQ ID NO: 3705:
US-09-107-433-3705

Query Match 100.0%; Score 3483; DB 15; Length 721;
Best Local Similarity 99.8%; Pred. No. 4,3e-289;
Matches 665; Conservative 1; Mismatches 0; Indels 0; Gaps 0:

QY 1 KIYNKNKLIADLGSERRVNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFLRN 60
Db 56 KIYNKNKLIADLGSERRVNAQANDIPDLVKAIVSIEDHFRFDRGIDTIRILGAFLRN 115
QY 61 LOSNSLOGSSTLTQOLIKLTYFSTSDOTS SRKQOEAFLAIQLEOKATKOEILYYINK 120
Db 116 LOSNSLOGSSTLTQOLIKLTYFSTSDOTS SRKQOEAFLAIQLEOKATKOEILYYINK 175
QY 121 VYMSNGNMGMTAAQNYGKDLNLSLPQALLAGMPAPNOYDPYSHPEAAQDRRLVL 180
Db 176 VYMSNGNMGMTAAQNYGKDLNLSLPQALLAGMPAPNOYDPYSHPEAAQDRRLVL 235
QY 181 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEVINQVEEETYNLLT 240
Db 236 SEMKNQGYISAEQYKAVNTPTDGLQSLKSASNPAYMDNYLKEVINQVEEETYNLLT 295
QY 241 TGMDDYTNVDOBAQKHLMDIYNTDEYVAYPPDELOVASTIVDSNGKYIAQIGARHOSN 300
Db 296 TGMDDYTNVDOBAQKHLMDIYNTDEYVAYPPDELOVASTIVDSNGKYIAQIGARHOSN 355
QY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNMNR 360
Db 356 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNMNR 415
QY 361 GYFGNITLQYALQOSRNPAYETLNKVGILNRKPTFLNGLGIDYPSIHYSNATSSNTTESD 420
Db 416 GYFGNITLQYALQOSRNPAYETLNKVGILNRKPTFLNGLGIDYPSIHYSNATSSNTTESD 475
QY 421 KKYGSSSKMAAAYAFANGGTYYPKMYIHKVYFSDGSEKESNNGTRAMKETTAYMMTD 480
Db 476 KKYGSSSKMAAAYAFANGGTYYPKMYIHKVYFSDGSEKESNNGTRAMKETTAYMMTD 535
QY 481 MKKTVLTGTRNAYLAMPQAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 540
Db 536 MKKTVLTGTRNAYLAMPQAGKTGTSNYTDEELENHIKTSQFVAPDELFAGYTRKYSM 595
QY 541 AVMTGYSNRLPLVNGNGLTVAAYKVRSMWTYISEGSPEDMNIPEGILRNGEFVRKNGAR 600
Db 596 AVMTGYSNRLPLVNGNGLTVAAYKVRSMWTYISEGSPEDMNIPEGILRNGEFVRKNGAR 655
QY 601 STWNSPAPQOPSTSSSSSSSDSSSTSPSTNNSTTTPNNNTTQOOSNTTPOQONN 660
|||||
```

Db 656 STWNSPAPQPPSTRESSSSSDSSTSSSTPTTNNSTTTNNNTTQOOSNTTPOOQON 715
Qy 661 POPAP 666
Db 716 POPAP 721

RESULT 6
PCT-US02-03987-13423
; Sequence 13423, Application PC/WUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitza Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits C
; FILE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; PRIOR FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US02-03987-13423

Query Match 99.5%; Score 3466; DB 1; Length 719;
Best Local Similarity 99.4%; Pred. No. 1.2e-287;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYDKRKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
Db 54 KIYDKRKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
Qy 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQEAALQLEOKAKOELITYYINK 120
Db 114 LOSNSLOGSALTLQOLIKLTYFSTSDQTSRKAQEAALQLEOKAKOELITYYINK 173
Qy 121 YMSNGNTGMOTAAONYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAODRRNLVL 180
Db 174 YMSNGNTGMOTAAONYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAODRRNLVL 233
Qy 181 SEMKNOGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNLYLKEVINOVEETGYNLLT 240
Db 234 SEMKNOGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNLYLKEVINOVEETGYNLLT 293
Qy 241 TGMDEVYTNVDEAOKHMDIYNTDEVYVAPDELOVASTIVDSNGKYIAOLGARHOSN 300
Db 294 TGMDEVYTNVDEAOKHMDIYNTDEVYVAPDELOVASTIVDSNGKYIAOLGARHOSN 353
Qy 301 VSEGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEBYNPGTNPVYNNMR 360
Db 354 VSEGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEBYNPGTNPVYNNMR 413
Qy 361 GYFGITITQYALQOOSRNPAYETLKNVGLNRAKPTFLNGIDYPSIHSNATSSMTTSSD 420
Db 414 GYFGITITQYALQOOSRNPAYETLKNVGLNRAKPTFLNGIDYPSIHSNATSSMTTSSD 473
Qy 421 KKYGASSEKMAAAYAFANGGTYYKPMYIHKYVFSDGSEKESNNGTAAKMTTAYMMTD 480
Db 474 KKYGASSEKMAAAYAFANGGTYYKPMYIHKYVFSDGSEKESNNGTAAKMTTAYMMTD 533
Qy 481 MKKTYLVTGGRNAYLANLPQAGKTGTSNYDEETENHIKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTYLVTGGRNAYLANLPQAGKTGTSNYDEETENHIKTSQFVAPDELFAGYTRKYSM 593
Qy 541 AVMTGYSRLLPVLVNGLTVAAYKVRSMYTYISEGSNPEDMNIIPGLRNGEFVKKNGAR 600
Db 594 AVMTGYSRLLPVLVNGLTVAAYKVRSMYTYISEGSNPEDMNIIPGLRNGEFVKKNGAR 653
Qy 601 STWNSPAPQPPSTRESSSSSDSSTSSSTPTTNNSTTTNNNTTQOOSNTTPOOQON 660

Db 654 STWSSPAPQPPSTRESSSSSDSSTSSSTPTTNNSTTTNNNTTQOOSNTTPOOQON 713
Qy 661 POPAP 666
Db 714 POPAP 719

RESULT 7
US-09-815-242-13423
; Sequence 13423, Application US/09815242

; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13423

Query Match 99.5%; Score 3466; DB 22; Length 719;
Best Local Similarity 99.4%; Pred. No. 1.2e-287;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KIYDKRKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
Db 54 KIYDKRKNLIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
Qy 61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAQEAALQLEOKAKOELITYYINK 120
Db 114 LOSNSLOGSALTLQOLIKLTYFSTSDQTSRKAQEAALQLEOKAKOELITYYINK 173
Qy 121 YMSNGNTGMOTAAONYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAODRRNLVL 180
Db 174 YMSNGNTGMOTAAONYGKDLNNLSLPOLALLAGMPAPNOYDPSHPEAODRRNLVL 233
Qy 181 SEMKNOGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNLYLKEVINOVEETGYNLLT 240
Db 234 SEMKNOGYISAOYEKAVNTPTDGLQSLKSASNPAYMDNLYLKEVINOVEETGYNLLT 293
Qy 241 TGMDEVYTNVDEAOKHMDIYNTDEVYVAPDELOVASTIVDSNGKYIAOLGARHOSN 300
Db 294 TGMDEVYTNVDEAOKHMDIYNTDEVYVAPDELOVASTIVDSNGKYIAOLGARHOSN 353
Qy 301 VSEGINQAVETNRDMSGTMKPTTDYAPALEYGVYDSTATIVHDEBYNPGTNPVYNNMR 360

```

|||||
Db 354 VSEGINQAVETNMDGSTMKPIITDYAPALEYGYVESTATIVHDEPYNGTNPVYNMNR 413
Qy 361 GYFGNTTLOYALQOOSRNVAVETLNKVGILNKRATPLNGIGIDIPSIHYSNAISSNTTESD 420
Db 414 GYFGNTTLOYALQOOSRNVAVETLNKVGILNKRATPLNGIGIDIPSIHYSNAISSNTTESD 473
Qy 421 KRYGASSEKMAAAYAAPANGGTYYKPMYIHKVYSDGSEKEFSNVGTTRAKETAYAMTD 480
Db 474 KRYGASSEKMAAAYAAPANGGTYYKPMYIHKVYSDGSEKEFSNVGTTRAKETAYAMTD 533
Qy 481 MKKTIVLYTGGRNAYLAMLPOAGKTGTSNYTDEIENHIKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTIVLYTGGRNAYLAMLPOAGKTGTSNYTDEIENHIKTSQFVAPDELFAGYTRKYSM 593
Qy 541 AVWTGYSNRLTPLYGNGLYVAAKVYRSMTYIYSEGSNPEDMNTPEGLYRNGEVEFRNGAR 600
Db 594 AVWTGYSNRLTPLYGNGLYVAAKVYRSMTYIYSEGSNPEDMNTPEGLYRNGEVEFRNGAR 653
Qy 601 STWNSPAPQOPPESTESSSSSDSTSSSTPSTNNSTTTNNNTTQOOSNTTPOOQON 660
Db 654 STWNSPAPQOPPESTESSSSSDSTSSSTPSTNNSTTTNNNTTQOOSNTTPOOQON 713
Qy 661 POPAP 666
Db 714 POPAP 719

```

RESULT 8

```

US-10-072-851-13423
; Sequence 13423, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: prf
; ORGANISM: Streptococcus pneumoniae
US-10-072-851-13423

```

```

Query Match 99.5%; Score 3466; DB 24; Length 719;
Best Local Similarity 99.4%; Pred. No. 1.2e-287;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
Db 54 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
Qy 61 LOSNSILOGSSTLTQOLIKLTYSTSTSDQTSKRKAQEWALIAOLEKATKOEILTYIYNK 120
Db 114 LOSNSILOGSSTLTQOLIKLTYSTSTSDQTSKRKAQEWALIAOLEKATKOEILTYIYNK 173

```

```

Qy 121 VYKSNNGYMGOTAAQNYVYGKDLNNLSLPOLALLAGNPOAPNOYDPSHPEAAQDRRLV 180
Db 174 VYKSNNGYMGOTAAQNYVYGKDLNNLSLPOLALLAGNPOAPNOYDPSHPEAAQDRRLV 233
Qy 181 SEMKNQYISAEOYERAVNTPITDGLQSLKSASNPAYMDNLYKEYINQVEETGYNLT 240
Db 234 SEMKNQYISAEOYERAVNTPITDGLQSLKSASNPAYMDNLYKEYINQVEETGYNLT 293
Qy 241 TGMDEVYTNDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDSNCKVIAOLGARQSSN 300
Db 294 TGMDEVYTNDOEAKHLMIDYNTDEVVAYPDDELQVASTIVDSNCKVIAOLGARQSSN 353
Qy 301 VSEGINQAVETNMDGSTMKPIITDYAPALEYGYVDSTATIVHDEPYNGTNPVYNMNR 360
Db 354 VSEGINQAVETNMDGSTMKPIITDYAPALEYGYVESTATIVHDEPYNGTNPVYNMNR 413
Qy 361 GYFGNTTLOYALQOOSRNVAVETLNKVGILNKRATPLNGIGIDIPSIHYSNAISSNTTESD 420
Db 414 GYFGNTTLOYALQOOSRNVAVETLNKVGILNKRATPLNGIGIDIPSIHYSNAISSNTTESD 473
Qy 421 KRYGASSEKMAAAYAAPANGGTYYKPMYIHKVYSDGSEKEFSNVGTTRAKETAYAMTD 480
Db 474 KRYGASSEKMAAAYAAPANGGTYYKPMYIHKVYSDGSEKEFSNVGTTRAKETAYAMTD 533
Qy 481 MKKTIVLYTGGRNAYLAMLPOAGKTGTSNYTDEIENHIKTSQFVAPDELFAGYTRKYSM 540
Db 534 MKKTIVLYTGGRNAYLAMLPOAGKTGTSNYTDEIENHIKTSQFVAPDELFAGYTRKYSM 593
Qy 541 AVWTGYSNRLTPLYGNGLYVAAKVYRSMTYIYSEGSNPEDMNTPEGLYRNGEVEFRNGAR 600
Db 594 AVWTGYSNRLTPLYGNGLYVAAKVYRSMTYIYSEGSNPEDMNTPEGLYRNGEVEFRNGAR 653
Qy 601 STWNSPAPQOPPESTESSSSSDSTSSSTPSTNNSTTTNNNTTQOOSNTTPOOQON 660
Db 654 STWNSPAPQOPPESTESSSSSDSTSSSTPSTNNSTTTNNNTTQOOSNTTPOOQON 713
Qy 661 POPAP 666
Db 714 POPAP 719

```

RESULT 9

```

US-09-107-532-5667
; Sequence 5667, Application US/09107532
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7308
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085598
; FILING DATE: May 14, 1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:

```

```

? NAME: Arianello, Pamela Deneke
? REGISTRATION NUMBER: 40,489
? REFERENCE/DOCKET NUMBER: CTC-012
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (781)893-5007
? TELEFAX: (781)893-8277
? INFORMATION FOR SEQ ID NO: 5667:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 823 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? HYPOTHETICAL: YES
? ORIGINAL SOURCE:
? ORGANISM: Enterococcus faecium
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 1...823
? OS-09-107-532-5667

```

Query Match	53.5%	Score 1862.5	DB 15	Length 823
Best Local Similarity	53.9%	Pred. No. 7e-150		
Matches 371	Conservative 109	Mismatches 179	Indels 29	Gaps 8

QY	1	KIYONKOLLADGSESRVANAANDIPDYLKVAI:IEHRFPNHR8IDPTRLILGALRN	60
Db	115	KLYIQOGELEPDLGAERREKISANELKTELEDIAIVSEDRFRYKHIGVDIRIIGSALSN	174
QY	61	LOSNSLOGGSTALLQOOLIKLAVFSTSTSDQTSIRKAOEAMLAIOLEOKATQOEILTYIINK	120
Db	175	FTSSGGLQSTLITQOOLIKLISFFSTSAEDQTLKRAQOEAAMAVRLEQKSKOEILTYIYNK	234
QY	121	VYMSNGVNGMOTAAONYNGDNLNLSIPOTALLAGMOPANOVDPRYSHREAADORRLYL	180
Db	235	VYMSNGLIGMETSAEMTFGKKLSBELSPOTALLAGMOPASAPDPRYVYRPOAKRRRDTYL	234
QY	181	SEMNQOYISAEOYEAENVPTITDGLQSLKSASNPYAPMNYLKEVINYQOEETGYILLT	240
Db	295	YTMLOÑEKISTQETDYQAVNPVPTDGLQELQOSDNDNTKIYDNYKEVINQOEKTDKNVYT	354
QY	241	TGMQVYTNVQEOAKHLMIDYINDEYAYVPEDELOVASTIVDYNSCKVYNQOLGARHOSSN	300
Db	355	DGLEIYTNLMDADACKRKLXYDINVDQYVYDPDEMQVASTLIDNTGCVKXQOIGRHAED	414
QY	301	VSEFINQAVEENBMOGSTMKPTIDYAPALEGYVDSTATVHDEPNYPGNTNPPVNMDR	360
Db	415	VTLGNNLNAVNTSRQFSGTMMKPVTDYGAPEFLKX - SIGKITTIDPARYBETSTIPVGMWDN	473
QY	361	GYFENITTLQALQOOSNNVPAVETLNKYGLNRKATFELNGCIDYPSIHYNSAISNTTESD	420
Db	474	QYMOTITLRLQALYLSRRVPAVKLFENEQSDKVASFLNLGIEVSTIHOSNAISSNTEEOD	533
QY	421	KKKVASSEKMAAAYAPAFANGCTYKRMVYIHKVYFSSQSGSKESPNVGTMRAMKEETVAMMT	479
Db	534	GTKKGASLKLMAAAYAPAFANGCTYKRYQYKNKIVFDQTEIEYEPDQTKTAMSEPTALMIT	553
QY	480	DMKTVTLTYGTGRNAYLAWLPQAKGTGTSNYTDEIEBNHAKTQSOFAVDELFAGYTRKYS	539
Db	594	DILKDTITTEGCTNAQIAGLYQAKGTGTSNYTDEYAK - LGISGVYPDILFAGYTPNYS	652
QY	540	MAYVATGYSNRILTPLYVGNGLVYAAKYVYSMMTYLSEGSNPFEDMNIPEBLR - NQEPVFKN	597
Db	653	ISVATGYNKKMTPTVTSSESHVASDYVELMQOYASANTYNTDMEPSELIVYGELLYKQD	712
QY	598	GARSTWNSPAPQOP-----PSTRSSSSSDSSTQSSSSTPTPNSTN-----	639
Db	713	YTAASNAITPSTTIPSSSYQTPESPSTTETTTQSSSSTQSESESTAESKSKSTTAETSEPA	772
QY	640	--TTPNANNTOQSTN-----TPDOONQNP	661
Db	773	SSTTVPSSSSSESSSTPESSAPPASSSEP	800

```

10 RESULT
11 US-09-107-532A-5667
12 : Sequence 5667, Application US/09107532A
13 : GENERAL INFORMATION:
14 : APPLICANT: Lynn A Doucette-Stamm and David Bush
15 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
16 : ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
17 : NUMBER OF SEQUENCES: 7310
18 : CORRESPONDENCE ADDRESSES:
19 : ADDRESSEE: GENOME THERAPEUTICS CORPORATION
20 : STREET: 100 Beaver Street
21 : CITY: Waltham
22 : STATE: Massachusetts
23 : COUNTRY: USA
24 : ZIP: 02354
25 : COMPUTER READABLE FORM:
26 : MEDIUM TYPE: CD/ROM ISO9660
27 : COMPUTER: PC
28 : OPERATING SYSTEM: <Unknown>
29 : SOFTWARE: ASCII
30 : CURRENT APPLICATION DATA:
31 : APPLICATION NUMBER: US/09/107,532A
32 : FILING DATE: 30-Jun-1998
33 : PRIOR APPLICATION DATA:
34 : APPLICATION NUMBER: 60/085,598
35 : FILING DATE: 14 May 1998
36 : APPLICATION NUMBER: 60/051571
37 : FILING DATE: July 2, 1997
38 : ATTORNEY/AGENT INFORMATION:
39 : NAME: Arinbello, Pamela Deneke
40 : REGISTRATION NUMBER: 40,489
41 : REFERENCE/DOCKET NUMBER: GTC-012
42 : TELECOMMUNICATION INFORMATION:
43 : TELEPHONE: (781)893-5007
44 : TELEFAX: (781)893-8277
45 : INFORMATION FOR SEQ. ID NO: 5667:
46 : SEQUENCE CHARACTERISTICS:
47 : LENGTH: 823 amino acids
48 : TYPE: amino acid
49 : TOPOLOGY: linear
50 : MOLECULE TYPE: protein
51 : HYPOTHETICAL: YES
52 : ORIGINAL SOURCE:
53 : ORGANISM: Enterococcus faecium
54 : FEATURE:
55 : NAME/KEY: misc.feature
56 : LOCATION: (B) LOCATION 1...823
57 : SEQUENCE DESCRIPTION: SEQ ID NO: 5667:
58 : US-09-107-532A-5667

```

Query Match	53.5%;	Score 1862.5;	DB 15;	Length 823;
Best Local Similarity	53.9%;	Pred. No. 7e-150;		
Matches 371;	Conservative 109;	Mismatches 179;	Indels . 29;	Gaps 8

Qy	1	KIINKNKNOIADGSESRVANAQDIPDLVYKALVSIEDHREFRGRIDIRILIGAPLN	60
		:	
Db	115	KLYIDGELFEDDLAERREKISANELPKTYLEDVAVSEDRFFKHIVDPRIIGSALSN	174
Qy	61	LOSNSLOGGATYTOOLIKLYTFSTSTSDOTISRKAEQAMWLIOLEQKATKOEILTYINK	120
Db	175	FTSGGLOGGATYTOOLIKLTFSTSTSDOTIKRAQAMAWAVRLEQKSKSOEILTYINK	234
Qy	121	VYMSNGYTGMOATAONTYYGKDNLNLSTPQALLAGMPQAPNOVDYPSHPERAODRRVLV	180
Db	235	VYMSNGYTGMPATAEMTFGKRISLSTLPQALLAGMPQAPSAVDPYVPPOAKRRRPVLV	294
Qy	181	SEMSNOCYISAEQYKAVNPFITGQSLKSSASNYPAYMONTYKEVINOVEEYGYLLT	240
Db	295	YTMLONEKISTEYDQAVANVPYVDGLODELQSDONTNTYIVDNYKEVINQYQEKTRKVVY	354
Qy	241	TGMVDVYTNVDEAKHLMDIYNTEDEVAVYPDEDLQVASTYVDVSNCKVIAQLGARHOSSN	300

```

Db 355 DGEIETNLDLAOKKLYIVMTDQYVSYPDEMOQVASTLIDINTGKVAQIGRHAIED 414
QY 301 VSHGICQAVETNRDMSGTMKPTIDYAPALEGYVDSSTATIVHDEPNYGTNPVYNNMR 360
Db 415 VTLGNLAVNTSDFOSTGKPVTDIGPAFEYLYK -STGKTIDAPYNEGTSTPVGNMD 473
QY 361 GYFNGTITLOALQOOSNRVAVETLNKVLNRAKTFPLNGLIDIPSIHYSNAISSNTTESD 420
Db 474 QYMGVITLQOALYSRNVPAVKLFNEVSGDKVASFLKNIGIEVSTIHOSNAISSNTEED 533
QY 421 -KRYGASSEKMAAAYAFANGCTYKPMYIHKVYSDGSEKESNVGTRAKETTYVMYT 479
Db 534 GTRYGASSTLKMAAAYAFANGCTYKPMYIHKVYSDGSEKESNVGTRAKETTYVMYT 593
QY 480 DMKKTLYTGTGNALVLAUPOAGKTGTSVYDELENIHKTSQFAPADELFGYTRKXS 539
Db 594 DILKDTITREGTGNAOIALYQNGKTGTSVYDELENIHKTSQFAPADELFGYTRKXS 652
QY 540 MAVTGYSNRLPPLVGNGLTVAAYKVRSMYTLSEGSNPEDMNIPGLYR-NGEYFVK- 597
Db 653 ISVMTGYNKKMPYVTESSSHVADYRELQYVSANVTMTDEMPSGLLRVGGELYKXQ 712
QY 598 -GARSTWNSPAPQP-----PSTESSSSSDSTSOSSSTPTSTNNST----- 639
Db 713 YPARSAIIPSTTIPSSSYVQTPGSSSTTETTTQSSSTQSESTASKESTTAETSEPA 772
QY 640 --TTNPNNTQOOSNT-----TPDOONONP 661
Db 773 SSTYVPSSSSESTPSSSAPPASSEP 800

```

RESULT 11
PCT-US02-03987-10728

```

; Sequence 10728, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitza Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for identifying the target of a Compound which inhibits
; FILE OF INVENTION: Proliferation
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; PCT-US02-03987-10728

```

Query Match 50.3%; Score 1751; DB 1; Length 778;
Best Local Similarity 51.4%; Pred. No. 2,4e-140;
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

```

QY 1 KIYDNKNOLADISERRRNAOANDIPTDLVKAIVSIEDHREPDHGGITIRILGAFLRN 60
Db 81 KIYDINNEFEEDGAEKRELIDPNDVPOLKDAIVSEDERRFKHGVDPTRITGSALSN 140
QY 61 LOSNSLOGSGSTLLOOLIKLTFSTSTSDOTISRKAQOEWAMLAIOLEOKATKOELLTYIYNK 120
Db 141 VANGGLOGSGSTLLOOLIKLTFSTSTSDOTISRKAQOEWAMLAIVLEBKSEELLTYIYNK 200
QY 121 VYMSNGNYMOTPAONYKYKDLNNLSLPOLALIGMPQAPNDYDPSHPEAADRNLVYL 180
Db 201 VYMANGFYEMETAENYKYKHLSELDLPOLALLAGMPQAPNSYDPTKPDATAKERDYYVL 260
QY 181 SEMKNGYISADYKAVMTPTIDGQSLKSASNYPAYMDNYLKEVINOYEETGNILTL 240
Db 261 YTMYDKRKISKAEYERAKATPIDEGLVPLKASDNRKVVYDNYKEVINEVAKTGNVYV 320

```

```

QY 241 TGMDDVYTNDOEAOKHLMDIYNTDEXVYAPPDDELQVASTIVDSNGKVIQAOLGARHOSSN 300
Db 321 DGLDIYTNLDMNAOKOLYDIVNSDQYVAFPPDDKMQVASFVIDVASGQVRAQIGRHAIPDD 380
QY 301 VSHGICQAVETNRDMSGTMKPTIDYAPALEGYVDSSTATIVHDEPNYGTNPVYNNMR 360
Db 381 VOLGNLAVNTQDVSSTYKPIMDYGPALLENLNY -STGRIMVDPKPKYRPTDIDVNSDI 439
QY 361 GYFNGTITLOALQOOSNRVAVETLNKVLNRAKTFPLNGLIDIPSIHYSNAISSNTTESD 420
Db 440 TYOGVITMRALINGSRNTAVQTFDEVKENIMPLFKGLGIDIKNLASNAISSNTEED 499
QY 421 -KRYGASSEKMAAAYAFANGCTYKPMYIHKVYSDGSEKESNVGTRAKETTYVMYT 479
Db 500 GDRYGISLKLAAAYAFANGCTYKPMYIHKVYSDGSEKESNVGTRAKETTYVMYT 559
QY 480 DMKKTLYTGTGNALVLAUPOAGKTGTSVYDELENIHKTSQFAPADELFGYTRKXS 539
Db 560 DMLKDVNLNGGTGNALVLAUPOAGKTGTSVYDELENIHKTSQFAPADELFGYTRKXS 619
QY 540 MAVTGYSNRLPPLVGNGLTVAAYKVRSMYTLSEGSNPEDMNIPGLYR-NGEYFVK- 599
Db 620 VSVMTGYNDRNTIYQYGIASDVRELMSYLSQVANSNDWVQPSYAVRNGELYKXA 679
QY 600 RSTWN-----SPAPQPSTESSSSSDSTSOSSSTPTSTNNSTTTNPNNTQOOSN 651
Db 680 YEVPYQVLEPSTSSAPQPESSSTVSSSTKEAESSSSSESAPASSEAPPESTEQPASS 739
QY 652 TTPDOONONPQAPQ 666
Db 740 SSAEQPATSEQPEP 754

```

RESULT 12

US-09-815-242-10728

```

; Sequence 10728, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Cair, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-815-242-10728

```

```
Query Match          50.3%; Score 1751; DB 22; Length 778;
Best Local Similarity 51.4%; Pred. No. 2.4e-140;
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KYIDKNOLADLGSERRNAQANDIPDYLKAIYSIEDHREDFHGRGIDTIRILGAFLRN 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 81 KYIDINNEIFEDLGAERKRLIQPNVDPOLKDAIVSVEDERRFYKHIGVDPIRILGALSIN 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 LOSNLSOGSSTLTQOLIKLTYFSTSDOTISRKAOEAMLAIQLEKATKOEILTYIYNK 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 141 YKNGLSOGSSTLTQOLIKLTYFSTSDOTISRKAOEAMLAIRLEKREKIEILTYIYNK 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 VYMSNGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAOPNOYDPYSHPAQAODRRNLVL 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 201 YVMANGFYGMETAEENYGYKHLSELDLPOTALLAGMPAOPNSYDPTKPTDKERRDVYL 260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 SEMKNOGYSABOYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLIT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 261 YTMNDKRIKSAEYKAKATPIDEGLVPLKASDNRKVVADNVYKEVINEVKAKTKGNVYT 320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 TGMDEVYTNVDOAKHMDIYNTDEYVAYPDELOVASTIYDVSNGKYIAOLGARHOSN 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 321 DGLDITLMDMAKQOLYDINSDOYVAFPPDKMQVASTVIDVAGQVRAQIGRHIPDD 380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 VSEGINOAETNRDMSGTMRKPTIDYAPALEYGYDSTATIVHDEPYNPGTNTPYVNMDR 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 381 VOLGNLNAVNTORDVGSYTKPIMDYGPALENLNT-STGLMLMDKPKRYKPGFTIDVFNSDL 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 GYFGNITLOVALQOSRNPVAVETLNKVLGNRAKTFNLGLGIDYPSIHSNALSNTTESD 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 440 TYOGVITMRRAIMGRNTJAVOTFDEVGKRNIMPFIKGLIDYKNLEASNALSNTSDVD 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 -KKYGASSEKMAAAVAAPANGSTYKPMYIHKVYFSDGSEKEFSVNGVTRAMKETAAYMNT 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 500 GDKYGISLKLAAVAAPANGSTYKPMYIHKVYFSDGSEKEFSVNGVTRAMKETAAYMNT 559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 DMKTVLTYGTGRNAYLAMLPOAGKTGTSNTDEIEENHIKTSQFVAPDELFGYTRKYS 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 560 DMLKDVNLNGGTFNGALIGLLOAAKTGTSNTDEDLARMGTTEKGIADPSTFVGTTTHYA 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 540 MAVMTGYSNRLPLVGNGLTYAAKYRSMNTYLSGSPNEDNINIEGLYRNGEYFEKNGA 599
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 620 VSWMTGYNDNRNPTIYQEYGYGASDYREIMSYLSQNVSDNDWQVDSVYRNGENELYKDA 679
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 600 RSTWN-----SPAOPPSTESSSSSDSTSSSTSTSTSTSTSTSTSTSTSTSTSTST 651
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 660 YEVPVQVLPSTTSSAPPESSSTYESSSTYKAESSSSSSSSSAPSSSEAPSTEQPASS 739
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 TTPDOONPOPAOP 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 740 SSAEQPATSEOPPEP 754
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-10-072-851-10728
; Sequence 10728, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
```

```
; FILE REFERENCE: ELITRA.028A
; CURRENT APPLICATION NUMBER: US/10/072.851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 10728
; LENGTH: 778
; TYPE: PRF
; ORGANISM: Enterococcus faecalis
US-10-072-851-10728

Query Match          50.3%; Score 1751; DB 24; Length 778;
Best Local Similarity 51.4%; Pred. No. 2.4e-140;
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KYIDKNOLADLGSERRNAQANDIPDYLKAIYSIEDHREDFHGRGIDTIRILGAFLRN 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 81 KYIDINNEIFEDLGAERKRLIQPNVDPOLKDAIVSVEDERRFYKHIGVDPIRILGALSIN 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 LOSNLSOGSSTLTQOLIKLTYFSTSDOTISRKAOEAMLAIQLEKATKOEILTYIYNK 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 141 YKNGLSOGSSTLTQOLIKLTYFSTSDOTISRKAOEAMLAIRLEKREKIEILTYIYNK 200
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 VYMSNGNTGMOTAAQNYGKDLNNLSLPOLALLAGMPAOPNOYDPYSHPAQAODRRNLVL 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 201 YVMANGFYGMETAEENYGYKHLSELDLPOTALLAGMPAOPNSYDPTKPTDKERRDVYL 260
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 SEMKNOGYSABOYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINOVEETGYNLIT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 261 YTMNDKRIKSAEYKAKATPIDEGLVPLKASDNRKVVADNVYKEVINEVKAKTKGNVYT 320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 TGMDEVYTNVDOAKHMDIYNTDEYVAYPDELOVASTIYDVSNGKYIAOLGARHOSN 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 321 DGLDITLMDMAKQOLYDINSDOYVAFPPDKMQVASTVIDVAGQVRAQIGRHIPDD 380
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 VSEGINOAETNRDMSGTMRKPTIDYAPALEYGYDSTATIVHDEPYNPGTNTPYVNMDR 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 381 VOLGNLNAVNTORDVGSYTKPIMDYGPALENLNT-STGLMLMDKPKRYKPGFTIDVFNSDL 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 GYFGNITLOVALQOSRNPVAVETLNKVLGNRAKTFNLGLGIDYPSIHSNALSNTTESD 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 440 TYOGVITMRRAIMGRNTJAVOTFDEVGKRNIMPFIKGLIDYKNLEASNALSNTSDVD 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 -KKYGASSEKMAAAVAAPANGSTYKPMYIHKVYFSDGSEKEFSVNGVTRAMKETAAYMNT 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 500 GDKYGISLKLAAVAAPANGSTYKPMYIHKVYFSDGSEKEFSVNGVTRAMKETAAYMNT 559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 DMKTVLTYGTGRNAYLAMLPOAGKTGTSNTDEIEENHIKTSQFVAPDELFGYTRKYS 539
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 560 DMLKDVNLNGGTFNGALIGLLOAAKTGTSNTDEDLARMGTTEKGIADPSTFVGTTTHYA 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 540 MAVMTGYSNRLPLVGNGLTYAAKYRSMNTYLSGSPNEDNINIEGLYRNGEYFEKNGA 599
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 620 VSWMTGYNDNRNPTIYQEYGYGASDYREIMSYLSQNVSDNDWQVDSVYRNGENELYKDA 679
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 600 RSTWN-----SPAOPPSTESSSSSDSTSSSTSTSTSTSTSTSTSTSTSTSTSTST 651
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 660 YEVPVQVLPSTTSSAPPESSSTYESSSTYKAESSSSSSSSSAPSSSEAPSTEQPASS 739
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 652 TTPDOONPOPAOP 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 740 SSAEQPATSEOPPEP 754
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-09-134-000-4939
; Sequence 4939, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; APPLICANT: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
```

;; TITLE OF INVENTION: FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: GTC-005
;; CURRENT APPLICATION NUMBER: US/09/134,000A
;; CURRENT FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 6810
;; SEQ ID NO 4939
;; LENGTH: 789
;; TYPE: PRT
;; ORGANISM: Enterococcus faecalis
US-09-134-000-4939

Query Match 50.3%; Score 1751; DB 15; Length 789;
Best Local Similarity 51.4%; Pred. No. 2.4e-140;
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;

QY 1 KIYDNKNOLIADLGSERRVNAOANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 60
DB 92 KIYDNKNOLIADLGSERRVNAOANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 151
QY 61 LOSNSLOGGSTLTLOOLIKLTFYESTSDOTISRKAOEAMLAIOLEOKATKOELITYYINK 120
DB 152 VANGSIGOGSTLTLOOLIKLTFYESTSDOTISRKAOEAMLAIVLEKSKKEILLITYYINK 211
QY 121 VYMSNGNYGMOTPAONYGYKDLNNLSLPOLALLAGMPQAPNOYDPSHPDEAADRRLV 180
DB 212 VYMANGFYEMETAENYGYKHSELDLPOTALLAGMPQAPNSYDPTKRDFAKERDYL 271
QY 181 SEMKNOGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLKEVINOVEETGYNLT 240
DB 272 YMYDNRKISKAEYERKAKTPIDEGLVPLKASDNRKVVNDYVKEVINERKAKTKRNVYT 331
QY 241 TGMADVNTVNDQEAOKHLMIDYNTDEYVAYPDELOVASTIVDSNGKVIAOLGARHOSN 300
DB 332 DGLDITYNLDMNAQOLYIVNSDOYVAFPPDKMOYASTIVDASQVQAOLGGRIIPDD 391
QY 301 VSEFGINQAVETNRDWSITMKPTTDYAPALEGYVDSTACIVHDEPYNPTGTNPPVYNNDR 360
DB 392 VOLGNLNAVNTORDVGSYVKKPIMDYGPALLENLNY-STGRILAVDKPKTKYPGTIDIVNSDL 450
QY 361 GFENGTITLOALQOSNRNVAVELINKVGLNRAKTPFLNGIGIOTPSHYHNAISSNTSED 420
DB 451 TYOGVITMRRAIAGSNTNTAYOTFDEVGKENIMPTFKGLGIDYKNLEANAISNSNSD 510
QY 421 -KKYGASSKMAAAAYAFANGSTYKPMYIHKVYFSDGSEKESFNNGTARMEETAYMMT 479
DB 511 GKTYGSSIKLAAAYAFANGSTYKPMYIHKVYFSDGSEKESFNNGTARMEETAYMMT 570
QY 480 DMKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEIEHNIKTQFVAPDELPAGYTRKYS 539
DB 571 DMKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEIEHNIKTQFVAPDELPAGYTRKYS 630
QY 540 MAVNMGYSRRLTPLYGNGLTVAAKYRSMYTLSEGSNEDMNIPBGLYRNGEFVYKNA 599
DB 631 VSMVNGYNDRNPTIYOEYGYIASDVYREIMSYLSQNSVNDVQPPSVYVAGNELVKA 690
QY 600 RSTWN-----SPAPQOPSTESSSSDSTSSSTSTSTSTSTSTSTSTSTSTSTSTSTST 651
DB 691 YEVPRVQVLPSTSSAPQOPSESSSYESSSTKRAESSSSSSSESAPSSSAPSTEDQPASS 750
QY 652 TTPDOONOPAPAP 666
DB 751 SSAEQPATSEQPEP 765

RESULT 15
US-08-116-541-4
; Sequence 4, Application US/08116541
; GENERAL INFORMATION:
; APPLICANT: Measure, H. Robert
; APPLICANT: Pearce, Barbara J.
; APPLICANT: Toumanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND

;; TITLE OF INVENTION: CELLULAR VACCINES BASED THERON
;; NUMBER OF SEQUENCES: 45
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Klaubert & Jackson
;; STREET: 411 Hackensack Avenue
;; CITY: Hackensack
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07601
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/116,541
;; FILING DATE: 19930901
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Jackson Esq., David A.
;; REGISTRATION NUMBER: 26,742
;; REFERENCE/DOCKET NUMBER: 600-1-069
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201 487-5800
;; TELEFAX: 201 343-1684
;; TELEX: 133521
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 320 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-116-541-4

Query Match 46.5%; Score 1621; DB 5; Length 320;
Best Local Similarity 99.7%; Pred. No. 7.2e-130;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADLGSERRVNAOANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 60
DB 5 KIYDNKNOLIADLGSERRVNAOANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLN 64
QY 61 LOSNSLOGGSTLTLOOLIKLTFYESTSDOTISRKAOEAMLAIOLEOKATKOELITYYINK 120
DB 65 LOSNSLOGGSTLTLOOLIKLTFYESTSDOTISRKAOEAMLAIOLEOKATKOELITYYINK 124
QY 121 VYMSNGNYGMOTPAONYGYKDLNNLSLPOLALLAGMPQAPNOYDPSHPDEAADRRLV 180
DB 125 VYMSNGNYGMOTPAONYGYKDLNNLSLPOLALLAGMPQAPNOYDPSHPDEAADRRLV 184
QY 181 SEMKNOGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLKEVINOVEETGYNLT 240
DB 185 SEMKNOGYISAOYEKAVNPTITDGLQSLKSASNPAYMDNLKEVINOVEETGYNLT 244
QY 241 TGMADVNTVNDQEAOKHLMIDYNTDEYVAYPDELOVASTIVDSNGKVIAOLGARHOSN 300
DB 245 TGMADVNTVNDQEAOKHLMIDYNTDEYVAYPDELOVASTIVDSNGKVIAOLGARHOSN 304
QY 301 VSEFGINQAVETNRDWS 316
DB 305 VSEFGINQAVETNRDWS 320

Search completed: June 13, 2002, 08:44:42
Job time: 248 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:40:34 ; Search time 20.69 Seconds
(Without alignments)
2575.142 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484
Sequence: 1 KIYDNKNQIADLGSRRVYN.....TQSNTPDQGNPQAP 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 244551 seqs, 79999522 residues

Total number of hits satisfying chosen parameters: 244551

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCN_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3484	100.0	666	4 US-08-961-083-2	Sequence 2, Appl
2	529.5	15.2	821	5 US-09-769-744A-122	Sequence 122, App
3	526	15.1	789	5 US-09-540-209B-8941	Sequence 8941, Ap
4	384	11.0	790	5 US-09-540-209B-9358	Sequence 9358, Ap
5	162	4.6	875	5 US-09-914-543-6	Sequence 6, Appl
6	160.5	4.6	2893	5 US-09-882-227-522	Sequence 522, App
7	159	4.6	615	5 US-09-882-227-408	Sequence 408, App
8	153.5	4.4	930	5 US-09-200-650D-3	Sequence 408, App
9	148	4.2	188	5 US-09-540-209B-6592	Sequence 6592, Ap
10	147.5	4.2	595	5 US-09-673-605A-34	Sequence 34, Appl
11	146	4.1	2843	5 US-09-442-489B-2	Sequence 2, Appl
12	143	4.1	1010	5 US-09-914-543-10	Sequence 10, Appl
13	141	4.0	2842	5 US-09-442-489B-7	Sequence 7, Appl
14	141	4.0	2843	5 US-09-987-482-1	Sequence 1, Appl
15	139.5	4.0	706	7 US-60-360-039-1865	Sequence 1865, Ap
16	134	3.8	994	5 US-09-935-625-7547	Sequence 7547, Ap
17	134	3.8	994	5 US-09-935-625-28205	Sequence 28205, A
18	134	3.8	1006	5 US-09-935-625-7546	Sequence 7546, Ap
19	134	3.8	1006	5 US-09-935-625-28204	Sequence 28204, A
20	133.5	3.8	478	5 US-09-445-289B-34	Sequence 34, Appl
21	133.5	3.8	541	7 US-60-360-039-9976	Sequence 9976, Ap
22	133.5	3.8	627	7 US-60-360-039-18489	Sequence 18489, A
23	132	3.8	1014	7 US-60-360-039-1608	Sequence 1608, Ap
24	131.5	3.8	1084	7 US-60-360-039-11155	Sequence 11155, A
25	131.5	3.8	1518	7 US-60-360-039-22243	Sequence 22243, A
26	131.5	3.8	10431	1 PCT-US02-11734-310	Sequence 310, App

27	131	3.8	989	5 US-09-935-625-7548	Sequence 7548, Ap
28	131	3.8	989	5 US-09-935-625-28206	Sequence 28206, A
29	131	3.8	1031	5 US-09-935-625-8784	Sequence 8784, Ap
30	131	3.8	1031	5 US-09-935-625-29529	Sequence 29529, A
31	131	3.8	1043	5 US-09-935-625-8783	Sequence 8783, Ap
32	131	3.8	1043	5 US-09-935-625-29528	Sequence 29528, A
33	129	3.7	708	5 US-09-540-209B-6575	Sequence 6575, Ap
34	128	3.7	1029	5 US-09-540-209B-10169	Sequence 10169, A
35	127.5	3.7	1609	7 US-60-360-039-1535	Sequence 1535, Ap
36	125	3.6	655	7 US-60-360-039-4713	Sequence 4713, Ap
37	125	3.6	655	7 US-60-360-039-7473	Sequence 7473, Ap
38	125	3.6	1166	5 US-09-200-650D-7	Sequence 7, Appl
39	124.5	3.6	1169	5 US-60-360-039-22288	Sequence 22288, A
40	124	3.6	650	7 US-60-360-039-22177	Sequence 22177, A
41	124	3.6	889	7 US-60-360-039-14305	Sequence 14305, A
42	124	3.6	889	7 US-60-360-039-14653	Sequence 14653, A
43	124	3.6	889	7 US-60-360-039-15126	Sequence 15126, A
44	123	3.5	1504	7 US-60-360-039-2471	Sequence 2471, Ap
45	122	3.5	648	5 US-09-540-209B-9400	Sequence 9400, Ap

ALIGNMENTS

RESULT 1
US-08-961-083-2
Sequence 2, Application US/08961083
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961.083
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P8340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-961-083-2
Query Match 100.0%; Score 3484; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 1.7e-217;
Matches 666; Conservative 0; Mismatches 0; Gaps 0;
OY 1 KIYDNKNQIADLGSRRVNAQNDIPDVKAIYSIDHRFPDHRGIDTIRIGAFARN 60

```

Db      1 KIYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRILGAFLRN 60
QY      61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAOEAMLAIOLEOKATKOELLTYINK 120
Db      61 LOSNSLOGSSTLTQOLIKLTYFSTSDQTSRKAOEAMLAIOLEOKATKOELLTYINK 120
QY      121 VYMSNGYMGQTAQNYKYKDLNLSLPOLALLAGMPQAPNOQDPYSHPEAADRNLVL 180
Db      121 VYMSNGYMGQTAQNYKYKDLNLSLPOLALLAGMPQAPNOQDPYSHPEAADRNLVL 180
QY      181 SEKNNGYISAEQYKAVNTPITDGLQSLKASANTPAYADNITKEVINOYEETGNLLT 240
Db      181 SEKNNGYISAEQYKAVNTPITDGLQSLKASANTPAYADNITKEVINOYEETGNLLT 240
QY      241 TCGDVTYNDQEAQKHLMTIYNTDEVYAPDDELQVASTIVDSNGKVIATOLGARQSSN 300
Db      241 TCGDVTYNDQEAQKHLMTIYNTDEVYAPDDELQVASTIVDSNGKVIATOLGARQSSN 300
QY      301 VSEGINQAVETNRDMSGTFMKPTDYAPALEYGYVDSSTATVHDEPNYPGTNPVYNNMR 360
Db      301 VSEGINQAVETNRDMSGTFMKPTDYAPALEYGYVDSSTATVHDEPNYPGTNPVYNNMR 360
QY      361 GFPGNTTLOALQOSRNVAVETLNKVGILNRAKTFNLGIDIPSIHYSNAISSNTTESD 420
Db      361 GFPGNTTLOALQOSRNVAVETLNKVGILNRAKTFNLGIDIPSIHYSNAISSNTTESD 420
QY      421 KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAKETATAMMD 480
Db      421 KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYFSDGSEKEFSNVGTRAKETATAMMD 480
QY      481 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDEIEHNHKTISOFAVADDELFAGYTRKYSM 540
Db      481 MKKTVLYTGGRNAYLAMLPOAGKTGTSNYTDEIEHNHKTISOFAVADDELFAGYTRKYSM 540
QY      541 AWTGGSNLTPLVGNGLVAAKYVRSMTYLSSEGSNPEDMNIPDLIYRNGEYFVNGAR 600
Db      541 AWTGGSNLTPLVGNGLVAAKYVRSMTYLSSEGSNPEDMNIPDLIYRNGEYFVNGAR 600
QY      601 STWNSAPQOPPESTESSSSSDSSTSTSTNNSTTTNNNTTQOQNTTPOOQON 660
Db      601 STWNSAPQOPPESTESSSSSDSSTSTSTNNSTTTNNNTTQOQNTTPOOQON 660
QY      661 POPAP 666
Db      661 POPAP 666

```

RESULT 2
US-09-769-744A-122
Sequence 122, Application US/09769744A

```

; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hamblift, Sean B
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PMC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 981636.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 122
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-122

```

```

Query Match      15.2%; Score 529.5; DB 5; Length 821;
Best Local Similarity 27.2%; Pred. No. 2.4e-26;
Matches 203; Conservative 115; Mismatches 270; Indels 159; Gaps 32;

QY      9 LTAOLGE--RRVNAQANDIPTDLVKAIVSIEDHREFDHDGIDTIRILGAFLRNLOS 63
Db      115 VIASIESDILRTSISEQISENKAIIATDEHFEHKGVPKAVIRATLTKEV--GLGS 173
QY      64 NSLOGSSTLTQOLIKLTYFSTSDQTSRKAOEAMLAIOLEOKATKOELLTYINKYVM 123
Db      174 SS--GGSTTLQOLIKQOYVDA---PTLARKAEIYDALALERAMKDELITTYLVAVAF 228
QY      124 SNGNYGM-----QTAQNYKYKDLNLSLPOLALLAGMPQAPNOQDPYSHPEAADRNLVL 169
Db      229 GRNNKQNTAGARQAAGIEFGVDASQUTVPOAFLAGLPSPITPYSPEYNTGELKSEDL 288
QY      170 EAAODRRNLVSEMKNOGYISAEQYKAVNTPITDGLQSLKASANTPAYADNITKEVINOYEETGNLLT 209
Db      289 EIGLRRAKAVLSMYRTGALSKDEYSQYKDYDLKODFLPSGVTGTISRDIYLYFTTLAEQ 348
QY      210 KSASNPAYMDNY-LKEVINQVE-----BETGYNLTGMDVYTNVDOEAQKH 256
Db      349 ERMYDTLQORDNVSAKELNKAETOKFYRLAKAEIENGCKITTT-----IDCKIHA 401
QY      257 LMDIYNTDEVYAPDD---ELQVASTIVDSNGKVIATOLGARQSSNVSEGINQAVETNR 313
Db      402 MQSA--VADGYLLDDGTGRTGVEGVNLMDNQTAIIGFVGGRNYQDNQ---NHADTRK 456
QY      314 DMSGTFMKPTDYAPALEYGYVDSSTATVHDEPNYPGTNPVYNNMRGTFGNTTLOALQ 373
Db      457 SPASTTKPLAAGIAIDQGLMS--ETILSNYPFNFGNPMIMANSKG--TGMMTGLDALN 514
QY      374 QSRNPAVER---LNKVGILNRAKTFNLGIDIPSIHYSNAISSNTTESDKRYGASSEK 430
Db      515 YSWNIRPAYMTYRLREKGD-VKGYMEKKGYELP-----EGIEISLPM 556
QY      431 AA-----AYAFANGGTYYKPMYIHKVYFSDGS-EKEFSNVGTRAKETATAMMD 478
Db      557 GCGIEVTAQHTNGYOTLANNGYVHOKHVISKEADGRVVEYQDKPQVYSKATATIM 616
QY      479 TDMKKTVLYTG-----TGRNAYLAMLPOAGKTGTSNYTDEIEHNHKTISOFAVADPE 529
Db      617 QGLLEVLSSRVTTTEKSNLTSLNPTLANADWIGKTGTIN-ODENWMLMLSTPRLTLGG- 674
QY      530 LFAGYTRKYSMAVWGTGYSN-----RLTPLV-GN-GLTVAAYKVSMTYLS 574
Db      675 -WIGHDHNLSLRAGYSNNNSNMAHLVNAIQOASPSIMCNERFALDPVYKSEV-LKST 732
QY      575 GSNPEDMNIP-----EGLYRNGEYFVFKNGARST-----WNSPAPQOP- 611
Db      733 GKPEKYSVGEKREVEVTGVTSTYMWKSGAPATYSRAIGSGADADYQUNAMSSIVGSLPT 792
QY      612 PSTESSSSSDSSTSSSTSTSTNNSTTTNNNTTQOQNTTPOOQON 668
Db      793 PSSSSSSSSSSSDSSTSTSTNNSTTTNNNTTQOQNTTPOOQON 660

```

RESULT 3
US-09-540-209B-8941
Sequence 8941, Application US/09540209B

```

; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 8941
; LENGTH: 789
; TYPE: PRT

```



```

; LENGTH: 615
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-882-227-408

```

Query Match	4.68;	Score 159;	DB 5;	Length 615;
Best Local Similarly	21.58;	Pred. No. 0.014;		
Matches 84;	Conservative 64;	Mismatches 125;	Indels 118;	Gaps 19

```

QY 187 GYIAEYOEKRAVNPIPIUDGLOSLSASANNPYAMDNLYKEVINOV---EEETGYNLLT--- 241
Db 190 GYVOKOEEDKIT---LTTGKKVEKS-----ODHLKAAQONKIRPGKRVSFNFIQNS 240
QY 242 -----GMDVYTNVVOEAKHLMIDYNTNDEYAVYEDDELQVASTYVDNSN---GKVIA 290
Db 241 YTEVERLIDGYEVYLSVPLKQIRELETLDDTK-----DKLKKELLVGIINKKSEIIS 294
QY 291 QLGAR-----HOSSNVSEFGINQAVETNRDWSITMKPTDYAPALEYGVYDSTAT 339
Db 295 LASSKRFPNPAIKTSDYESLNLS-----VAEKVEEPGSTIKPIV-YSLDDKMLINPKER 348
QY 340 IVHDEPNYPOSTNPVYNNMDRGV--GNITL-----QYALQOSRNPVAVETLN 385
Db 349 I-----DLNHGYQLQKRYTIKODFIPSKRAVVEDILIOSSNGMT--- 388
QY 386 KVGJLN-RAKTFELNGL-----GIDYP-----SIHYSNAISSNTTESDKRYG---AS 426
Db 389 KISKNNLNKDFYNNLLGTYGSOKTGDIDLSLEATGKIPPLSAFRREVLYKGSVGYGLNAT 448
QY 427 SEKMAAAYAAFAFGTGYKKPMYIHKVVEFSDG-----SEKEFSNVGTRAKKETAYM 477
Db 449 FLQLLRAYAAVFNSECKLTTEPYLVQRETPANBGDIYIPSPKPTFOVIPSRSARKMKET--- 504
QY 478 MTDMAKTYLVYTGGRNAYLMLPQAGKTGTS 508
Db 505 ---LIVVRGYTGKNAQFGLYIGTGTGA 531

```

```

RESULT      8
US-09-200-650D-3
; Sequence 3, Application US/09200650D
; GENERAL INFORMATION:
; APPLICANT: Foster, Timothy J.
; APPLICANT: Eichlin, Deirdre N.
; APPLICANT: Hook, Magnus A.O.
; APPLICANT: Perkins, Samuel L.
; TITLE OF INVENTION: Extracellular Matrix-Binding Proteins from Staphylococcus aureus
; FILE REFERENCE: P06283u9/BA8
; CURRENT APPLICATION NUMBER: US/09/200,650D
; CURRENT FILING DATE: 1998-11-25
; PRIOR APPLICATION NUMBER: 60/066,815
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: 60/098,427
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: patentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-200-650D-3

```

Query Match	4.4%	Score 153.5	DB 5	Length 930
Best Local Similarity	20.4%	Pred. No. 0.058		
Matches 155	Conservative 96	Mismatches 288	Indels 221	Gaps 37
QY	59	RNLQNSLGGSGTTLQOLIKLYTSTYSDDTISRKAEAWLAIDLEQAKTKOELLFYFI	118	
Db	80	ROLKDNLT-----QTAAADQPKYIMDSATVKTETSSN-----MSPONATINQSTTKTS	127	
QY	119	NKVY--MSNGNYGMQTAQNTY--YGRDLNNL-----SLPOLLV--LAGMPQAPNQYD	164	

Db	128	NVTTNDKSSFTTYSNEEDKSNLTQAKDVSTTPKTTTIKPRTLNMAANFVAAPQOGTNVND	187
Qy	165	P--SYHPEAAOORRNILVLESEMKNOCY-ISAEOYERKANVPITDGLD	207
Db	188	KHFSNIDLAIDKGVHNOQTGTKEFWATSSDVLKLANITIDSVAGEGDTFFKQOYR	247
Qy	208	--SLKSASN---YAPAYMDNYLKEVINQVEETGVLLTTGMADVTVNV-----	249
Db	248	PCSVRLRPSOTOLYNXNOGMIIAKGIYDSTTNNTTPT-FTNYVDQOYNNVKSQFEOVAFNR	306
Qy	250	-----DOEQOKHLMIDYNNFDEVAVYPPDELOQVASTIVDVSNGCVIAOQARHSSNVSG	304
Db	307	KNAITDKTYAKKEVITLGN-DTY-----SEEL-----IVDYGKKAQPLISNTNTINNEDLS	356
Qy	305	INQAVETNNDGWSYTKMP-----ITDY---APALEYGVYDSTATIVDEPNYVPGTNPVY	356
Db	357	RMTAVVNOCKMTYTQTQTFVTNLGTGKFNPNNAKFRIVEY-----DQNOFVDSFP--	408
Qy	357	NMDRGYFGGITTQY-ALQOSRNPVAVETLKNVGLNAAKIFL-----NLGID	402
Db	409	--DTSKLMQVTDQFDVIYANDKNTATVJDLMKGGTSSNKQYIIQOAVYPDMSSTDNG-KID	465
Qy	403	YP-----SIHVSNAISSNTTESD--KKYVASSEKMAAAVAAPANGTY-----	443
Db	466	YLLDDTKTYSMSNSISNVNGSSSTANGDOKY-----NLGDIWEDTNKD	510
Qy	444	-----YKPMYIHKVYVSDGSEKEFSNVGTRANKETTYAVMMTDMKTVLTYGTGRNA	494
Db	511	GKODANERKIGKGY---VILKDSNGKELDRTTT--DENGKQYFTGLSNG--TYVSEFT	562
Qy	495	YLAWLPOAKCTGTSNTTDEIENHIKTSQVAPD---ELFAGY--TRKYSAAVMWGY--	546
Db	563	PAGYPTPTANVG---TDDAVYSDGLTGTGVJKDADNMJLDSGFYTKPRKSLIDAYWYDS	618
Qy	547	-----SNRLPFLYGNGLVVAAYKVRSMWTVYSEGSNPEPDWMINPEG	586
Db	619	NNDGRKDSITEKIGKVKYVILQNEKEGVIIGTETLEDENGKR--FDNLDGKYYKVIIFEPKAG	676
Qy	587	LYRNG-----BEVFKNG--ANSTWNSPAPQOPPSTRESSSSSS	621
Db	677	LQGTGNTTDEDDKADGGEVDVTLTIDHDPFTLNDNGYEETSDSDSDSDSDSDSDSDSDS	736
Qy	622	DS-SFSQSSSTPPTNNSTTTNPNNTTQOOSNTTPDOONQ	660
Db	737	DSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDSDS	776

```

RESULT      9
US-09-540-209B-6592
; Sequence 6592, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 6592
; LENGTH: 188
; TYPE: PRT
; ORGANISM: B.fragilis
US-09-540-209B-6592

```

```
Query Match Similarity      4.2%   Score 148; DB 5; Length 188;
Best Local Similarity     27.6%   Pred. No. 0.014;
Matches    48; Conservative 34; Mismatches 76; Indels 16; Gaps 6;

Qy      38 EDHREDFNRHGIDTIRLGAFLRNLO-SNSIAGSSTLTUOLIKLYFSTSDOTSRRKAO 96
          ||::|||::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      4 EDGAFVYHNGEFLPEPMRKALYODLKVRFRARSGSTTMOLVSVEFS--RKNLNARKLE 60
```


[illegible]

```

RESULT 12
US-09-914-543-10
: Sequence 10, Application US/09914543
: GENERAL INFORMATION:
: APPLICANT: DIVERSA CORPORATION
: APPLICANT: IAM, David
: APPLICANT: MATHUR, Eric
: TITLE OF INVENTION: ENDOGLUCANASES
: FILE REFERENCE: DIVER1150-5
: CURRENT APPLICATION NUMBER: US/09/914,543
: CURRENT FILING DATE: 1998-11-22
: PRIOR APPLICATION NUMBER: PCT/US97/08793
: PRIOR FILING DATE: 1997-05-22
: PRIOR APPLICATION NUMBER: US 08/651,572
: PRIOR FILING DATE: 1996-05-22
: NUMBER OF SEQ ID NOS: 52
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10
: LENGTH: 1010
: TYPE: PRF
: ORGANISM: Tereidinbacter
US-09-914-543-10

```

Query Match 4.1%; Score 143; DB 5; length 1010;
 Best Local Similarity 19.5%; Pred. No. 0.31;
 Matches 131; Conservative 111; Mismatches 285; Indels 146; Gaps 29

[illegible]

```

Db 265 DGAPAVMEIQMMDFELKÖNNISHTLNMVSVDKLBGASIVQ-----Pg--TPISGM----- 312
Qy 364 GNITLÖYALAOOSRNVPAVEFLNK-VGINRAKTEFLNGLIDYPSIHTHSMASISFTTSDDK 422
Db 313 -----ASBLTASGLTVKNIYVNMGTJIGNG-----SSSSSSSSSSSSSSSS 355
Qy 423 YGASSEKMAAAYAFPAANGGTYYKPMYTHKVVFS---DGSEKESNYTRAMKFTTAYMMT 479
Db 356 SSSSSSSSSSSSSSGTGGNCAGVNYPMWTARDMSGALYNNHANGDÖWYONSLYLRAN 415
Qy 480 DMKRTVLYTGGRNAYLAWLPÖAGAKTTSNVYDEIEHNHAKTSOFAPADELFAGYTRKYS 539
Db 416 WYTNKSV----PGDASMTSLGACGCGNSTSSSSSSSSSSSSSSSSSGTGGSSSSSS 471
Qy 540 MAVMTGYSNRLTPLVGNGLVAAYKVRSMNTYLSSEGSNPEDMNIPEGLYRNGEYFKNG- 598
Db 472 SSSSSSSSSSSSGGCGCTEVCWYQOGTYPLCJNNNTSGWMENNOISCIÖRCESSONG 531
Qy 599 -----ASTWMSPAQÖPPTRESSSSSDBSTSQSSSTPTSTNNSTJTNNNNTQOOSNTJ 653
Db 532 AGGVVSNCTGSSSTSS---SSSSSSSSSSSSSSSSSSSSSGTSSSTSSSSSSSSSSST 588
Qy 654 PDÖÖNONPOPAÖP 666
Db 589 GSSGMPGPRVDNP 601

```

```

RESULT 13
US-09-442-489b-7
Sequence 7, Application US/09442489B
GENERAL INFORMATION:
APPLICANT: Albertsen, Hans
APPLICANT: Anand, Rakesh
APPLICANT: Carlson, Mary
APPLICANT: Groden, Joanna
APPLICANT: Hedger, Philip John
APPLICANT: Joshi, Geoff
APPLICANT: Kintzer, Kenneth
APPLICANT: Markham, Alexander Fred
APPLICANT: Nakamura, Yusuke
APPLICANT: Thliveris, Andrew
APPLICANT: Vogelstein, Bert
APPLICANT: White, Raymond L.
TITLE OR INVENTION: APC Antibodies
FILE REFERENCE: 001107.78817
CURRENT APPLICATION NUMBER: US/09/442.489B
CURRENT FILING DATE: 1999-11-18
PRIOR APPLICATION NUMBER: US 08/452,654
PRIOR FILING DATE: 1995-05-25
PRIOR APPLICATION NUMBER: US 08/289,548
PRIOR FILING DATE: 1994-08-12
PRIOR APPLICATION NUMBER: US 07/741,940
PRIOR FILING DATE: 1991-08-08
NUMBER OF SEQ ID NOS: 154
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2842
TYPE: PRT
ORGANISM: Homo sapiens
US-09-442-489b-7

```

Query Match	4.0%:	Score 141;	DB 5;	Length 2842;
Best Local Similarity	18.7%:	Pred. No. 1.8;		
Matches 129;	Conservative 99;	Mismatches 213;	Indels 250;	Gaps 31

```

QY 56 ALENNLOSN-----LGGSTLQOLIKLFTSTSDQTSRKAGEAMLAIDLE 106
      |||||:::||||:
Db 723 AALNLANPRAKKDANIMSPGSSLSLHVR-----KQ 756
      |||||:::||||:
QY 107 KATKOELLTYINKVWNSNGY-----GMQTAQNYGKIDNNISLPOLALLAGMPQA 159
      ||:::||||:
Db 757 KALPAELDAOHLSTFPNNIDLSPKASHRSKORHKSLYDY-----799
      ||:::||||:

```

[illegible]

Query Match	4.0%	Score 141;	DB 5;	Length 2843;
Best Local Similarity	18.7%	Pred. No. 1.8;		
Matches 129;	Conservative 99;	Mismatches 213;	Indels 250;	Gaps 31.

Oy	56	AFLRLNQSSNS-----	LOGGRTLTQQLIKLTFSTSTSPQTSRKQEAMLAQLEQ	106
		: :		
Db	724	AALRLNLMANRPARYKDANIMSPGSLPSLHVR-----		KQ 757
Oy	107	KATKOELITLYINKRYMNSGN-----	GMQTAQNYGYGRDLNNLSLPQALLALGMPPQA	159
		: : : : :		
Db	758	KALELELAQHLSTFFNDLNDLSPKASHRSKQRHKOISLYGDVV-----		800

```

QY 160 PNQDYPHPPAADRRN-----LYLSEMKNGIISADQYERKAVNPPIIDGLOSLKSASNY 215
Db 801 ---EPTNRHSDNRSDNNTGNMTVLSY-----LMTTV-----LPSSSS 837
QY 216 PAYMDNYLKEVINOVEEETGNLLTTGMDVYTNVDQEAOKHLMADYINDEYVAPDDE-- 273
Db 838 RGSIDSSSEMDRLEBERG-----IGLGNV-----HATENP 870
QY 274 -----LQVASTIVDVSNGKVIAGLCARHOSSNVSFGINQAVETNRDNGST--MKPIYD 324
Db 871 GTSSKRGQIISTAAQIA--KVMEEVSAHTS-----QEDRSSGSTTELHCYTD 917
QY 325 YAPALEYVYDSTATIYHDEPYNPGNTPYVWMDGYESGNTILOXA--LQOSRNPVAVET 383
Db 918 ERNALR-----RSSAHTHSNTYN-----TKSENSNR--TCSMPRAKLEYRRS--SDWS 963
QY 384 LNKVGLN-----RAKTFPLNGLID-----YPS-----IHYSAISNTTESDK 421
Db 964 LNSVSSSGGYCKRQCKMPISLESSEDESKFCYGYRADLANHKIHSANHMNDNGELDT 1023
QY 422 -----KYGASSEKMAAAYAFANGGTYYKPMYIHKVVEPSDSEKEFESVNGTRAKETTA 475
Db 1024 PLYNSLKY--SDEQLNGRGQSPQSNBERMAPKHIIDEIKOSEQRQSRN-----QSTT 1074
QY 476 YMDMDMKATVLYTGTGNATLANLPPQAGKTGTSNTYDELEHNIKTSGVAFADFLFAGYT 535
Db 1075 YPV-----YTESDNDHHLKFPQHPGOOEVSYPYR 1103
QY 536 RKYSAVWGTGYSNRLPYLGNGLTVLAAKYRSM-----LYLSEGSNPEMDNPIRE 585
Db 1104 SR-----GANGSETNRKVSNNHGIMONVVSQSLCOEDDYEDOKPINYSEPRYSSEEPHHEE 1166
QY 586 GLYRNGEVEFKNGARISWNSP-----AFQOPPTSESSSSSDSTSOSSSTT--PSTNN 637
Db 1157 ERPTVYSIKYNEEKRRHY--DQPIDYSLKYADPIDPSQKOSFSPKSSSGSSKTEHMSSS 1215
QY 638 STTTPNNNTQOSN-----TTPDOONQNPORA 664
Db 1216 ENTSTPSSNAKRONQILPPSSAQSNSGQPOKA 1246

```

```

RESULT 15
US-60-360-039-1865
; Sequence 1865, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1865
; LENGTH: 706
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-60-360-039-1865

```

Query Match	Similarity	4.0%	Score	139.5	DB	7	Length	706	
Best Local	Similarity	19.0%	Pred	NO	0.32				
Matches	146	Conservative	115	Mismatches	221	Indels	279	Gaps	40
QY	58	LRN-LQSNLSGCGSTLTOOLIKRYEFTSTG	-----DQTSRK-----	AQ	96				
Db	3	LRNIIOASSLILGSLIAAD	-----SSSTGEGYAPSIIPCPSDDTSLVRNAGSLTAE	55					
QY	97	EAMLAIOLEQKATKOEILLYINKRYMSGNGYGMQTAQNYGCKLNNLSLPOL	-----	150					

```
Db      56 TDWLK---KRDAYTKREALHSLR---ATSNFSDTSLSTLFSS--NSSNPKIGIACSG 107
QY      151 ----ALLAGMPQAPNOYDPSHPEAODRNLVLSEMNQGIYSAQYEKAVN----- 199
Db      108 GGYRAMLGAGACMI-----RANDKRIDGANEHGLIGLSSSTYISGLSGNMLTG 156
QY      200 TPITDGLSLKASNYPAYMD---NYLKEVINQVEETGYNLLTGMVYTNVDQEAQ-- 254
Db      157 TLAMNNMTSVQEIYDHMSESDSNMTITSIVN----PGSNLYTITIERMESIYQVQAK 211
QY      255 -----KHLMD---IYNDEYVAYPDDELQVA-STIVDY--SNGKV--IAQLGAR 295
Db      212 SPAGFNISLSDLMARALSYNF--FPLSDPAGSALTWSSSLRDVDVFKNGEMPLPIIVADCR 269
QY      296 HSSSNVSGICINAVETNMDGSTMKPIIDYAPALFEGYVDSTATIVHDEPYNPGTNT-- 353
Db      270 YPGTIV-INLNT-----LFEFTF-FEMGSWDPSLNAFTD--VKYIGTNVTN 312
QY      354 --PYNMDRGYFGNITLQYALQOSRNV-----PAVETLNKVGILNRAKTFNLGLGIDYP 404
Db      313 GKPV-NKQOCVSGYDNAGVITASLNFESLEASTSTYTKMINSFANKYVNNLSQDD 371
QY      405 SI-----HYSNA----- 411
Db      372 DIAIYAANPFKDFEVDNRNYTSSIVDADDLFLVGDGEDGQNLPIVPLIKKERDLVVFAL 431
QY      412 -ISSNTTES-----DKKYGASSEKMAAFA---AFANGTYKKPMYT----- 449
Db      432 DISDNTDESMPGVCMTNTERKQYKQKGMAFPYPVPVNTFNLGLTNKPTFFGCDAKN 491
QY      450 -----HKVFSDSSEKEFS-NVGTIRAMKETTAYMMTDMKTVLTYGTGRNAYLAWLP 500
Db      492 LTLDLFIPIPLVYIIPNTHSFENGOSTLKN---YNVTERL-----GMIRNGF----- 536
QY      501 QAGKTGTSNYTDE-----IENHKTISQFVAPDELFAGYTRKYSMAVMTGYSNRL 550
Db      537 EAATMG--NFTDSDNPLGICIGCAIIRKQESLNATLPRPC---TKCFADYCMNG----- 585
QY      551 TPLVNGGLTVAKAYRSMMTYISEGSNPEDMNIPGLIRNGEFVFKNGARSTWNSPAPQO 610
Db      586 -----TLSTANPE-----LSGNSTYQSGAIAASAISEATDG 616
QY      611 PP-----STESSSSSDSTSQSSSTPSTNNSTTTNPNNTQOSN 651
Db      617 IPTALLGSSSTSGNTTSTSTSTSSNVTSSNSNSSNTTLNNSSSSS 663
```

Search completed: June 13, 2002, 08:41:39
Job time: 65 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:41:59 ; Search time 34.55 Seconds
(without alignments)
2141.104 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666

Sequence: 1 KIDYNNKQIADLGSERRVN.....TOOSNTTPOQNONPOPAQP 666

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 10

Total number of hits satisfying chosen parameters: 30

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A.Geneseq_032802:*

```
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	666	19	AAW55063
2	330	49.5	682	17	AAW04359
3	330	49.5	719	22	AAU37830
4	245	36.8	320	16	AAU70153
5	206	30.9	420	20	AAV56106
6	14	2.1	778	22	AAU35135
7	11	1.7	266	19	AAV86004
8	11	1.7	821	21	AAV81757
9	10	1.5	110	19	AAW55113
10	10	1.5	119	22	AAW69173
11	11	1.5	284	20	AAV49227

12	10	1.5	284	20	AAV32100
13	10	1.5	406	20	AAV49225
14	10	1.5	406	20	AAV32098
15	10	1.5	428	20	AAV49238
16	10	1.5	428	20	AAV32110
17	10	1.5	438	22	AAV92949
18	10	1.5	446	20	AAV49140
19	10	1.5	446	20	AAV49143
20	10	1.5	459	20	AAV49251
21	10	1.5	459	20	AAV32190
22	10	1.5	460	20	AAV48250
23	10	1.5	460	20	AAV32189
24	10	1.5	462	22	AAU03646
25	10	1.5	655	20	AAV49226
26	10	1.5	655	20	AAV32099
27	10	1.5	694	21	AAV81653
28	10	1.5	774	22	AAU36453
29	10	1.5	781	22	AAU35684
30	10	1.5	1704	15	AAV49657

ALIGNMENTS

```
RESULT 1
AAW55063
ID AAW55063 standard; Protein; 666 AA.
XX
AC AAW55063:
XX
DT 02-OCT-1998 (first entry)
XX
DE Streptococcus pneumoniae SP001 protein.
XX
KW Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
XX detection; pneumonia; otitis media; meningitis.
XX
OS Streptococcus pneumoniae.
XX
PN W09818930-A2.
XX
PD 07-MAY-1998.
XX
PF 30-OCT-1997; 97WO-US19422.
XX
PR 31-OCT-1996; 96US-0029960.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Chol GH, Hromocky J A, Johnson LS, Kunsch CA;
XX WPI: 1998-272224/24.
XX DR N-PSDB; AAV27323.
XX
PT Nucleic acid encoding antigenic peptide(s) from Streptococcus
XX pneumoniae - or their epitope-containing fragments, useful in
XX protective or therapeutic vaccines, and for diagnosis
XX
PS Claim 11; Page 48; 118pp; English.
XX
XX
XX The present sequence represents a protein from Streptococcus pneumoniae.
XX The nucleic acid sequence encoding the Streptococcus pneumoniae protein
XX can be useful in vaccines for inducing protective antibodies against
XX Streptococcus pneumoniae, for treatment or prevention of infection e.g.
XX pneumonia, otitis media or meningitis. Probes based on the nucleic acid
XX are used to detect Streptococcus infection (by usual hybridisation or
XX amplification methods), also for isolating Streptococcus genes or their
XX allelic variants. The protein can be used similarly to detect specific
XX antibodies in standard immunoassays, especially for diagnosing or
XX monitoring infections. Antibodies which bind the protein and for passive
XX detect corresponding antigens, to purify the protein and for passive
XX immunisation (optionally coupled to a toxin). Vaccines are administered,
XX e.g. by injection, orally or through the skin, typically at 0.01-1000
```

CC (especially 10-300) mu g/ml per dose.
 XX
 SQ Sequence 666 AA:

Query Match 100.0%; Score 666; DB 19; Length 666;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNOLADGSESRVNAQANDIPTDLVKAIVSTEDHREFDHQIDIRITIGAFLRN 60
 DB 1 KIYDNKNQIADIGSESRVNAQANDIPTDLVKAIVSTEDHREFDHQIDIRITIGAFLRN 60
 QY 61 LOSNSLQSGSTLTQOLIKLTFFSTSDQTISSRAQEWALAIQLEOKATKOEILTYIYINK 120
 DB 61 LQNSLQSGSTLTQOLIKLTFFSTSDQTISSRAQEWALAIQLEOKATKOEILTYIYINK 120
 QY 121 VYMSNGYNGQTAQNTYGGDLNNLSLPOLALLAGMPQANQYDPTSHPEAADRRLVL 180
 DB 121 VYMSNGYNGQTAQNTYGGDLNNLSLPOLALLAGMPQANQYDPTSHPEAADRRLVL 180
 QY 181 SEKNNGYISAEQYERAVNTPITDGLQSLKASNPAYMNYLKEVINQVEEETGYNLTT 240
 DB 181 SEKNNGYISAEQYERAVNTPITDGLQSLKASNPAYMNYLKEVINQVEEETGYNLTT 240
 QY 241 TGMDEVYTNDOEAQKHLMDIYNTDEVYAYPDDQLQVASTIVDSNGKVIQAQARRQSSN 300
 DB 241 TGMDEVYTNDOEAQKHLMDIYNTDEVYAYPDDQLQVASTIVDSNGKVIQAQARRQSSN 300
 QY 301 VSGINQAVETNRDMSGSTMKPIYDAPALFEGYDSTATTVHDEPNYPGTNTPYVWMDR 360
 DB 301 VSGINQAVETNRDMSGSTMKPIYDAPALFEGYDSTATTVHDEPNYPGTNTPYVWMDR 360
 QY 361 GYFENITLQVALQOOSRNVAPETLNFVGLNRKATFLNGLGIDIDYPSIHYSNAISSNTTESD 420
 DB 361 GYFENITLQVALQOOSRNVAPETLNFVGLNRKATFLNGLGIDIDYPSIHYSNAISSNTTESD 420
 QY 421 KKYGASSEKMAAAYAFANGTYKKPMYTHKVPFSDGSEKFEFNVGTRANKETAYMMTD 480
 DB 421 KKYGASSEKMAAAYAFANGTYKKPMYTHKVPFSDGSEKFEFNVGTRANKETAYMMTD 480
 QY 481 MMFTVLYTGRNAYLAWLQOAGKTGTSNNTDEIENHITISQFVAPDELFACTRRYSM 540
 DB 481 MMFTVLYTGRNAYLAWLQOAGKTGTSNNTDEIENHITISQFVAPDELFACTRRYSM 540
 QY 541 AAVTGTGSMRLTPLVGNGLTFAAKVYSMMYLTSEGSNPEDMNIPEGLYRNGEEVFKNGAR 600
 DB 541 AAVTGTGSMRLTPLVGNGLTFAAKVYSMMYLTSEGSNPEDMNIPEGLYRNGEEVFKNGAR 600
 QY 601 STWNSPAPQOPESTESSSSSDSSTSPSTNNSTTTPNNTTQOSNTTPDOONON 660
 DB 601 STWNSPAPQOPESTESSSSSDSSTSPSTNNSTTTPNNTTQOSNTTPDOONON 660
 QY 661 POPAOP 666
 DB 661 POPAOP 666

RESULT 2
 ID AAM04359 standard; protein; 682 AA.
 XX AAM04359;
 AC AAM04359;
 XX
 XX
 DT 03-DEC-1996 (first entry)
 XX
 DE S. pneumoniae penicillin binding protein 1A soluble variant.
 XX
 KW Penicillin binding protein; PBP 1A; bifunctional protein; transglycosylase; transpeptidase; identification; assay; inhibitor; antibiotic resistant; bacteria; soluble variant; protein structure; X-ray crystallography; determination.
 XX

OS Streptococcus pneumoniae.
 XX
 XX GB2290792-A.
 XX
 XX
 PD 10-JAN-1996.
 XX
 XX 29-JUN-1995; 95GB-0013306.
 XX
 XX 24-NOV-1994; 94SE-0004072.
 PR 01-JUL-1994; 94IN-0000580.
 XX
 PA (ASTR) ASTRA AB.
 XX
 PI Balganesch TS, Town CM;
 DR WPT; 1996-042232/05.
 DR N-FSDB; AAT08027.
 XX
 PT Sol. derivs. of bifunctional penicillin binding protein (BPBP) -
 PT opt. lack transglycosylase activity, useful to identify and assay
 PT for antibodies or cpds. which bind BPBP
 XX
 PS Claim 2; Pages 65-68; 108pp; English.
 XX
 CC The present sequence is a soluble variant of the S. pneumoniae
 CC penicillin binding protein (PBP) 1A, which lacks the 1st 38 amino
 CC acid residues of the wild type protein. Wild type PBP is a
 CC bifunctional protein, which binds the cell membrane when expressed
 CC in a bacterial cell, having transglycosylase and transpeptidase
 CC activities. The variant protein (NCIMB 40665) in conjunction with
 CC a labelled anti-bifunctional PBP monoclonal antibody, can be used
 CC to identify and assay for cpds. which bind bifunctional PBP. Such
 CC cpds., as inhibitors of bifunctional PBP have a potential use in
 CC therapeutic cpds. which inhibit the growth of antibiotic resistant
 CC bacteria. The soluble variant may also be used in X-ray
 CC crystallography.
 XX
 SQ Sequence 682 AA:
 Query Match 49.5%; Score 330; DB 17; Length 682;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 72 LFOOLIKLTFFSTSDQTISSRAQEWALAIQLEOKATKOEILTYIYINKYMSNGYNGMQ 131
 DB 88 LFGQLIKLTFFSTSDQTISSRAQEWALAIQLEOKATKOEILTYIYINKYMSNGYNGMQ 147
 QY 132 TAAQNTYNGKDLNNLSLPOLALLAGMPQANQYDPTSHPEAADRRLVLSEKNGYTISA 191
 DB 148 TAAQNTYNGKDLNNLSLPOLALLAGMPQANQYDPTSHPEAADRRLVLSEKNGYTISA 207
 QY 192 EOYERAVNTPITDGLQSLKASNPAYMNYLKEVINQVEEETGYNLTTGMDVYTNVDQ 251
 DB 208 EOYERAVNTPITDGLQSLKASNPAYMNYLKEVINQVEEETGYNLTTGMDVYTNVDQ 267
 QY 252 EAQKHLMDIYNTDEVYAYPDDQLQVASTIVDSNGKVIQAQARRQSSNVSFQINQAVET 311
 DB 268 EAQKHLMDIYNTDEVYAYPDDQLQVASTIVDSNGKVIQAQARRQSSNVSFQINQAVET 327
 QY 312 NRDWGSTMKRITDYAPALEYGVDSSTATTIVHDEPNYPGTNTPYVWMDRKYFGNITLQYA 371
 DB 328 NRDWGSTMKRITDYAPALEYGVDSSTATTIVHDEPNYPGTNTPYVWMDRKYFGNITLQYA 387
 QY 372 LOOSRNVPAVETLNFVGLNRKATFLNGLGIDYPSIHYSNAISSNTTPESDCKKYGASSEKMA 431
 DB 388 LOOSRNVPAVETLNFVGLNRKATFLNGLGIDYPSIHYSNAISSNTTPESDCKKYGASSEKMA 447
 QY 432 AAYAFANGSTYKKPMYTHKVPFSDGSEKFEFNVGTRANKETAYMMTDMMKTVLTGTG 491
 DB 448 AAYAFANGSTYKKPMYTHKVPFSDGSEKFEFNVGTRANKETAYMMTDMMKTVLTGTG 507
 QY 492 RNAVYLAWLQOAGKTGTSNNTDEIENHITISQFVAPDELFACTRRYSMAVMTGYSNRLT 551

XX Novel gene fragments encoding specific bacterial exported proteins
 PT - specifically of *S. pneumoniae*, useful as vaccines
 XX
 PS Claim 35; Page 88-9; 168pp; English.
 CC This sequence represents exp2. The DNA encoding this sequence is
 CC identical to that for pona which encodes penicillin-binding protein 1A
 CC (Pbp1a). This sequence is involved in adhesion of bacteria to target
 CC cells. This sequence is an exported protein of *S. pneumoniae*. Export
 CC proteins are the proteins in pathogenic bacteria that are virulence
 CC determinants. Other export proteins include p1pa (see AAK70152), exp1,
 CC exp3, and pad1 (encoded by the sequence shown in AAK83259). This
 CC sequence can be inserted into an expression vector (preferably a
 CC bacterial expression vector) to provide for high levels of expression of
 CC the protein. The protein can then be used in the production of an
 CC acellular vaccine. These vaccines are used to provide protection from
 CC Gram positive bacterial infection. Antibodies against export proteins
 CC can be used for diagnosis of infection and in passive immune therapy.
 CC
 SQ Sequence 320 AA;
 Query Match 36.8%; Score 245; DB 16; Length 320;
 Best Local Similarity 100.0%; Pred. No. 1e-233;
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 72 LTGOLIKLTFSTSPDITSRKAOEAWLAIOLKATKOELLTYINKVYMSNGYGMQ 131
 Db 76 lTgqLiKlLyStsdsqLstsrkaqaawlaIqLqkaTqellTyInkvmsngnygmq 135
 QY 132 TAAQNYVYKDLNNLSLPOLALAGMPQAPNQDYPYSHPEAODRRLVLEMKNOGYISA 191
 Db 136 taagnyyKdlnlslpqlallagmpqapnqdyprshpeaagdrnlvlekmngyisa 195
 QY 132 EGYEAVNTPITDGLQSLKASASNPAYMDNLYKEVINOVEETGYNLLTTGMDVYTNVDQ 251
 Db 136 egeYekavntptldglqslksasnypaymdnlykEvinoveetgynlltgmDytnvdq 255
 QY 252 EAKOHLMDITNDEVVAYPDDDELQVASTVDVNSGVINOLGARHOSNVSGINQAVET 311
 Db 256 eeqkhlmditndevvaypddelqvastlvdvnsngvlnolgarhgsnvsfignqavet 315
 QY 312 NRDMG 316
 Db 316 nrDMG 320
 RESULT 5
 AAY56106
 ID AAY56106 standard; Protein; 420 AA.
 AC AAY56106;
 XX 07-FEB-2000 (first entry)
 DE Streptococcus pneumoniae pbp1a TER isolate a) protein sequence.
 XX
 KW Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1a;
 KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
 KW detection; identification; pneumococcal meningitis.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN ZA9807024-A.
 XX
 PD 28-APR-1999.
 XX
 PF 05-AUG-1998; 98ZA-0007024.
 XX
 PR 01-AUG-1997; 97ZA-0006886.
 XX
 PA (SAME-) SOUTH AFRICAN INST MEDICAL RES.

PA (UWI-) UNIV WITWATERSRAND.
 XX (MEDI-) MEDICAL RES COUNCIL.
 XX
 PI Klugman KP, Smith AM, Du Plessis M;
 XX
 DR WPI, 1999-601770/51.
 DR N-PSDB; AA635939.
 XX
 PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
 CC useful for the diagnosis of pneumococcal meningitis
 PS Claim 11; Fig 4; 63pp; English.
 CC
 CC A polymerase chain reaction (PCR) assays have been developed for
 CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
 CC using primers based on the penicillin binding protein 2B (pbp2B) gene
 CC and the pbp1a gene. The products and methods can be used for detecting
 CC *S. pneumoniae*, particularly antibiotic-resistant strains. They can be
 CC used for simultaneously diagnosing pneumococcal meningitis and
 CC identifying any antibiotic-resistant *S. pneumoniae* strains in a sample.
 CC The methods can be used for detecting *S. pneumoniae* strains resistant
 CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
 CC The assays can be adapted to detect other pathogens causing meningitis.
 CC The assays can be used to detect an antibiotic resistant strain of
 CC *S. pneumoniae* with a minimum inhibitory concentration (MIC) of
 CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
 CC a 224 bp product. The present sequence represents a Streptococcus
 CC pneumoniae pbp1a transpeptidase encoding region (TER) isolate protein
 CC sequence from the present invention.
 CC
 SQ Sequence 420 AA;
 Query Match 30.9%; Score 206; DB 20; Length 420;
 Best Local Similarity 99.5%; Pred. No. 5.4e-195;
 Matches 406; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 196 KAVNPEITDGLQSLKASASNPAYMDNLYKEVINOVEETGYNLLTTGMDVYTNVDQ 255
 Db 5 kavnpeitdglqslksasnypaymdnlykEvinoveetgynlltgmDytnvdq 255
 QY 256 HLMIDITNDEVVAYPDDDELQVASTVDVNSGVINOLGARHOSNVSGINQAVETNRDM 315
 Db 65 hlmditndevvaypddelqvastlvdvnsngvlnolgarhgsnvsfignqavetnrDM 315
 QY 316 GSTMKRPTDYPALLEYGVYDSTATIVHDEPYNPGNFPYVNMDBCFGNITLOVALQOS 375
 Db 125 gstmkrptdyPALLEYGVYDSTATIVHDEPYNPGNFPYVNMDBCFGNITLOVALQOS 375
 QY 376 RNVPVAVETLNKGLNRAKTFNLGLIDYPSIHYSNALISSNTTESDCKYGASSEKMAAAYA 435
 Db 185 rnvpvavetlnkglNRAKTFNLGLIDYPSIHYSNALISSNTTESDCKYGASSEKMAAAYA 435
 QY 436 AFANGSTYKPMYIHKVVEVSDGSEKESNVGTRAKMETTAYMWTMDMKTVLTYTGGRNAV 495
 Db 245 afangstlykpmYIHKVVEVSDGSEKESNVGTRAKMETTAYMWTMDMKTVLTYTGGRNAV 495
 QY 496 LAMLPOAGKTGTSNTYDDELINHIKTSOFVAPDELFACTRTKYSMAVWVGYSNRLLTPVVG 555
 Db 305 lamlpOAGKTGTSNTYDDELINHIKTSOFVAPDELFACTRTKYSMAVWVGYSNRLLTPVVG 555
 QY 556 NGLTVAKVYRSMNTYLSGSSNPEDWNIPEGLYRNGEFPFKNGARSTW 603
 Db 365 ngltvaKVYRSMNTYLSGSSNPEDWNIPEGLYRNGEFPFKNGARSTW 603
 RESULT 6
 AAU35135
 ID AAU35135 standard; Protein; 778 AA.
 AC AAU35135;
 XX
 DT 13-FEB-2002 (first entry)


```

XX DE Enterococcus faecalis cellular proliferation protein #422.
XX
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX
XX OS Enterococcus faecalis.
XX
XX PN WO200170955-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 21-MAR-2001; 2001WO-US09180.
XX
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX
XX PA (ELIT-) ELITRA PHARM INC.
XX
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX
XX DR WPI: 2001-611495/70.
XX DR N-PSDB; AAS52994.
XX
XX PT New polynucleotides for the identification and development of
XX PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX PS Example 3; Seq ID No 10728; 511pp; English.
XX
XX CC The invention relates to antisense inhibitors of genes essential to
XX CC prokaryotic cellular proliferation, their use in identifying the
XX CC genes, their use in the discovery of novel antibiotics, the essential
XX CC genes themselves and the encoded proteins. The prokaryotes used are
XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
XX CC invention is also useful for the identification of potential new targets
XX CC for antibiotic development. The antisense nucleic acids can also be used
XX CC to identify proteins used in proliferation, to express these proteins,
XX CC and to obtain antibodies capable of binding to the expressed proteins.
XX CC The proteins can be used to screen compounds in rational drug discovery
XX CC programmes. The antisense nucleic acid sequence is also useful to screen
XX CC for homologous nucleic acids which are required for cell proliferation in
XX CC a wide variety of organisms. The present sequence represents an
XX CC essential prokaryotic cellular proliferation protein.
XX CC Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 778 AA:

```

```

Query Match 2.1%; Score 14; DB 22; Length 778;
Best Local Similarity 100.0%; Pred. No. 0.0001;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 66 LOGSTIQQLIKL 79
DB 146 Iggstltqqlikl 159

```

```

RESULT 7
AAV86004
ID AAV86004 standard; Protein; 266 AA.
XX AC AAV86004;
XX

```

```

DT 10-APR-2000 (first entry)
XX
XX DE S. pneumoniae derived protein #213.
XX
XX KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
XX KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.
XX
XX OS Streptococcus pneumoniae.
XX
XX PN WO9806734-A1.
XX
XX PD 19-FEB-1998.
XX
XX PF 15-AUG-1997; 97WO-US14436.
XX
XX PR 16-AUG-1996; 96US-0024022.
XX
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
XX PI Stodola RK;
XX
XX DR WPI: 1998-159452/14.
XX DR N-PSDB; AA296344.
XX
XX PT Streptococcus pneumoniae proteins and related DNA - useful for
XX PT screening compounds for antibacterial activity
XX
XX PS Claim 5; Page 493-494; 640pp; English.
XX
XX CC This invention describes novel isolated Streptococcus pneumoniae
XX CC polynucleotides (see AAV96173-796494) and their encoded proteins (see
XX CC AAV85792-186182). The DNA, vectors and host cells described in the
XX CC method of the invention are useful for the recombinant expression of the
XX CC polypeptides. The polypeptides are useful for treatment or prevention of
XX CC disease, or diagnosis of disease related to expression or activity of
XX CC such a polypeptide. They can also be used to screen for compounds which
XX CC interact with and inhibit or activate such a polypeptide. The
XX CC polypeptides (or DNA encoding them, via gene therapy) are also useful
XX CC for inducing an immunological response in a mammal. The antagonists are
XX CC useful to inhibit such bacterial polypeptides. The polypeptides are
XX CC particularly useful to identify antimicrobial compounds and antibiotics.
XX CC They are also useful to determine their role in pathogenesis of
XX CC infection, dysfunction and disease.
XX
XX SQ Sequence 266 AA:

```

```

Query Match 1.7%; Score 11; DB 19; Length 266;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 68 GGSITLQQLIK 78
DB 107 ggstltqqlik 117

```

```

RESULT 8
AAV81757
ID AAV81757 standard; Protein; 821 AA.
XX AC AAV81757;
XX
XX DT 02-JUN-2000 (first entry)
XX
XX DE Streptococcus pneumoniae protein sequence ID113.
XX
XX KW Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
XX KW bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
XX KW kidney disease; diabetes; immunosuppressive disorder; otitis media;
XX KW pneumococcal septicaemia; sinusitis; meningitis; therapy.
XX
XX OS Streptococcus pneumoniae.

```

```

XX PN WO200006738-A2.
XX XX
XX PD 10-FEB-2000.
XX XX
XX PF 27-JUL-1999; 99WO-GB02452.
XX XX
XX PR 27-JUL-1998; 98GB-0016336.
XX PR 19-MAR-1999; 99US-0125329.
XX XX
XX PA (MICR-) MICROBIAL TECHNIQS LTD.
XX PI Le Page RWF, Wells JM, Hanniffy SB, Hansbro PM;
XX XX
XX DR WPI: 2000-195301/17.
XX DR N-PSDB; AA291853.
XX XX
XX PT Streptococcal proteins and polynucleotides useful for diagnosis,
XX XX treatment and prophylaxis of bacterial infections
XX PS Claim 2: Page 59; 76pp; English.
XX XX
XX CC This sequence represents a Streptococcus pneumoniae protein of the
XX CC invention. The proteins (or their homologues, derivatives and/or
XX CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
XX CC compositions comprising the proteins are useful as vaccines and also in
XX CC diagnostic assays. The sequences are useful for the detection or
XX CC diagnosis of S. pneumoniae infection, by contacting a sample to be tested
XX CC with them. Agents capable of antagonising, inhibiting or interfering with
XX CC the function or expression of the protein or prophylactide are useful in
XX CC medical compositions in the treatment or prophylaxis of S. pneumoniae
XX CC infection. As the sequences can be used to treat S. pneumoniae infection,
XX CC they can be used to treat bacterial pneumonia, which has high rates in
XX CC young children, the elderly, and in patients with predisposing conditions
XX CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
XX CC or with immunosuppressive disorders, especially AIDS. They can also be
XX CC used to treat pneumococcal septicemia, otitis media, sinusitis, and
XX CC meningitis.
XX SQ Sequence 821 AA;

Query Match 1.7%; Score 11; DB 21; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSTFQOLIK 78
DB 176 ggstltqqlik 186

RESULT 9
AAW55113
ID AAW55113 standard; Protein; 110 AA.
XX AAW55113;
XX AC
XX XX
XX DT 02-OCT-1998 (first entry)
XX XX
XX DE Streptococcus pneumoniae SP0062 protein.
XX XX
XX KW Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
XX KW detection; pneumonia; otitis media; meningitis.
XX XX
XX OS Streptococcus pneumoniae.
XX XX
XX PN WO9818930-A2.
XX PI
XX PD 07-MAY-1998.
XX XX
XX PF 30-OCT-1997; 97WO-US19422.
XX PR
XX PR 31-OCT-1996; 96US-0029960.
XX PT

```

```

XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI
XX PI Choi GH, Hromocky J A, Johnson LS, Kunsch CA;
XX XX
XX DR WPI: 1998-272224/24.
XX DR N-PSDB; AAV27374.
XX XX
XX PT Nucleic acid encoding antigenic peptide(s) from Streptococcus
XX PT pneumoniae - or their epitope-containing fragments, useful in
XX PT protective or therapeutic vaccines, and for diagnosis
XX PS Claim 11; Page 70; 118pp; English.
XX XX
XX CC The present sequence represents a protein from Streptococcus pneumoniae.
XX CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein
XX CC can be useful in vaccines for inducing protective antibodies against
XX CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.
XX CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
XX CC are used to detect Streptococcus infection (by usual hybridisation or
XX CC amplification methods), also for isolating Streptococcus genes or their
XX CC allelic variants. The protein can be used similarly to detect specific
XX CC antibodies in standard immunoassays, especially for diagnosing or
XX CC monitoring infections. Antibodies which bind the protein are used to
XX CC detect corresponding antigens, to purify the protein and for passive
XX CC immunisation (optionally coupled to a toxin). Vaccines are administered,
XX CC e.g. by injection, orally or through the skin, typically at 0.01-1000
XX CC (especially 10-300) mu g/ml per dose.
XX SQ Sequence 110 AA;

Query Match 1.5%; Score 10; DB 19; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 18 sssssdst 27

RESULT 10
ABB69173
ID ABB69173 standard; Protein; 119 AA.
XX ABB69173;
XX AC
XX XX
XX DT 26-MAR-2002 (first entry)
XX XX
XX DE Drosophila melanogaster polypeptide SEQ ID NO 34311.
XX XX
XX KW Drosophila: developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX XX
XX OS Drosophila melanogaster.
XX XX
XX PN WO200171042-A2.
XX PD
XX PD 27-SEP-2001.
XX XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX XX
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX XX
XX PA (PEKE ) PE CORP NY.
XX PI
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX XX
XX DR WPI: 2001-656860/75.
XX DR N-PSDB; ABL13276.
XX XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more

```

PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure: SEQ ID NO 34311; 21pp + Sequence Listing; English.
 CC
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL016175) and the encoded proteins
 CC sequences (ABB57737-ABB72072).
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 119 AA;
 Query Match 1.5%; Score 10; DB 22; Length 119;
 Best Local Similarity 100.0%; Pred. No. 0.17; Mismatches 0; Gaps 0;
 Matches 10; Conservative 0; Indels 0;
 QY 617 SSSSSDSTS 626
 |||||
 DB 60 sssssdsts 69
 RESULT 11
 AAY49227
 ID AAY49227 standard; Protein: 284 AA.
 XX
 AC AAY49227;
 XX
 DT 07-FEB-2000 (first entry)
 XX
 DE N-terminal truncate of Cbpa serotype type 4 polypeptide R1.
 XX
 KW Choline binding protein A; Cbpa; truncate; immune response; infection;
 KW pneumococcal bacterium; vaccine.
 XX
 OS Streptococcus pneumoniae.
 OS
 PN WO9951187-A2.
 XX
 PD 14-OCT-1999.
 XX
 PF 07-APR-1999; 99WO-US07668.
 XX
 PR 07-APR-1998; 98US-0056019.
 PR 07-APR-1998; 98US-0080878.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 PI Tuomanen EI, Masure HR;
 XX
 DR WPI; 1999-620161/53.
 DR N-PSDB; AAZ31403.
 XX
 PT Novel polypeptides, used to develop products for the diagnosis,
 PT prevention and treatment of pneumococcal infections -
 XX
 PS Claim 3; Page 11; 85pp; English.
 XX
 CC The invention provides novel isolated polypeptides comprising the amino
 CC acid sequence of an N-terminal choline binding protein A (Cbpa) truncate
 CC (CbpaR). The polypeptides can be selected from sequences shown in
 CC AAY49225, AAY49227, AAY49230, AAY49231, AAY49233 and AAY49238. The
 CC polypeptides can be used for inducing an immune response in a subject
 CC which has been exposed to or infected with a pneumococcal bacterium.
 CC The polypeptides can also be used for preventing infection by a pneumococcal bacterium.
 CC Vaccines comprising the polypeptides or encoding nucleic acids can be

CC used for treating a subject infected with or exposed to a pneumococcal
 CC bacterium. Antibodies specifically binding the polypeptides can be used
 CC for detection and diagnosis and for preventing pneumococcal attachment to
 CC a mucosal surface. The products can be used in humans and other animals
 CC such as domestic animals, such as feline or canine subjects, farm animals
 CC such as bovine, equine, caprine, ovine, and porcine subjects, wild
 CC animals (whether in the wild or in a zoological garden), research
 CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
 CC i.e. for veterinary medical use.
 XX
 SQ Sequence 284 AA;
 Query Match 1.5%; Score 10; DB 20; Length 284;
 Best Local Similarity 100.0%; Pred. No. 0.37; Mismatches 0; Gaps 0;
 Matches 10; Conservative 0; Indels 0;
 QY 616 SSSSSDSTS 625
 |||||
 DB 113 sssssdst 122
 RESULT 12
 AAY32100
 ID AAY32100 standard; Protein: 284 AA.
 XX
 AC AAY32100;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE Choline binding protein A (Cbpa) truncate R1.
 XX
 KW Choline binding protein; Cbpa; truncate; adhesion; immunogen;
 KW vaccine; diagnosis; therapy; pneumococcus; sepsis; meningitis;
 KW otitis media; pneumonia.
 XX
 OS Streptococcus pneumoniae.
 OS
 PN WO9951188-A2.
 XX
 PD 14-OCT-1999.
 XX
 PF 07-APR-1999; 99WO-US07669.
 XX
 PR 07-APR-1998; 98US-0056019.
 PR 07-APR-1998; 98US-0080878.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA (MEDI-) MEDIMUNE INC.
 XX
 PI Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;
 XX
 DR WPI; 1999-633690/54.
 DR N-PSDB; AAZ34453.
 XX
 PT New N-terminal choline binding protein A truncate polypeptides, used to
 PT develop products for the diagnosis, prevention and treatment of
 PT pneumococcal infections -
 XX
 PS Claim 3; Page 93-95; 160pp; English.
 XX
 CC The present sequence represents polypeptide R1, a new N-terminal
 CC choline binding protein A (Cbpa) truncate derived from the Cbpa of
 CC Streptococcus pneumoniae serotype type 4. Host-vector systems for
 CC production of R1 and other Cbpa truncate polypeptides are provided.
 CC The invention relates generally to novel N-terminal Cbpa truncates
 CC and polynucleotides encoding them, host-vector systems, and
 CC antibodies that specifically bind to the polypeptides. The
 CC invention also relates to vaccines including Cbpa polypeptides,
 CC which provide protection or elicit protective antibodies to
 CC bacterial infection, specifically pneumococcus, and to use of
 CC antibodies and antagonists against such polypeptides in diagnosis

Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
 |||||
DB 113 ssssssdst 122

RESULT 15

AAV49238
ID AAV49238 standard; Protein; 428 AA.

AC AAV49238;

DT 07-FEB-2000 (first entry)

DE Choline binding protein A (Cbpa) serotype type 4 amino acid sequence.

KW Choline binding protein A; Cbpa; truncate; immune response; infection;
KW pneumococcal bacterium; vaccine.

OS Streptococcus pneumoniae.

PN W09951187-A2.

PD 14-OCT-1999.

PF 07-APR-1999; 99WO-US07668.

PR 07-APR-1998; 98US-0056019.

PR 07-APR-1998; 98US-0080878.

PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PI Tuomanen EI, Masure HR;

DR WPI; 1999-620161/53.

PT Novel polypeptides, used to develop products for the diagnosis,
PT prevention and treatment of pneumococcal infections -

PS Claim 7; Page 7; 85pp; English.

CC The invention provides novel isolated polypeptides comprising the amino
CC acid sequence of an N-terminal choline binding protein A (Cbpa) truncate
CC (CbpaT). The polypeptides can be selected from sequences shown in
CC AAV49225, AAV49227, AAV49230, AAV49231, AAV49233 and AAV49238. The
CC polypeptides can be used for inducing an immune response in a subject
CC which has been exposed to or infected with a pneumococcal bacterium. They
CC can also be used for preventing infection by a pneumococcal bacterium.
CC Vaccines comprising the polypeptides or encoding nucleic acids can be
CC used for treating a subject infected with or exposed to a pneumococcal
CC bacterium. Antibodies specifically binding the polypeptides can be used
CC for detection and diagnosis and for preventing pneumococcal attachment to
CC a mucosal surface. The products can be used in humans and other animals
CC such as domestic animals, such as feline or canine subjects, farm animals
CC such as bovine, equine, caprine, ovine, and porcine subjects, wild
CC animals (whether in the wild or in a zoological garden), research
CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
CC 1.e. for veterinary medical use.

SQ Sequence 428 AA;

Query Match 1.5%; Score 10; DB 20; Length 428;

Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
 |||||
DB 113 ssssssdst 122

RESULT 16
AAV32110
ID AAV32110 standard; Protein; 428 AA.

AC AAV32110;

DT 01-FEB-2000 (first entry)

DE Choline binding protein A (Cbpa) N-terminal truncate.

KW Choline binding protein; Cbpa; truncate; adhesin; immunogen;
KW vaccine; diagnosis; therapy; pneumococcus; sepsis; meningitis;
KW otitis media; pneumonia.

OS Streptococcus pneumoniae.

PN W09951188-A2.

PD 14-OCT-1999.

PF 07-APR-1999; 99WO-US07669.

PR 07-APR-1998; 98US-0056019.

PR 07-APR-1998; 98US-0080878.

PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
(MEDI-) MEDIMONE INC.

PI Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;

DR WPI; 1999-633690/54.

PT New N-terminal choline binding protein A truncate polypeptides, used to
PT develop products for the diagnosis, prevention and treatment of
PT pneumococcal infections -

PS Claim 7; Page 119-122; 160pp; English.

CC The present sequence represents a new N-terminal choline binding
CC protein A (Cbpa) truncate that maintains the tertiary structure of
CC streptococcus pneumoniae serotype type 4 Cbpa native protein.
CC Host-vector systems for production of N-terminal Cbpa truncates are
CC provided. The invention relates generally to novel N-terminal
CC Cbpa truncates and polynucleotides encoding them, host-vector
CC systems, and antibodies that specifically bind to the truncates.
CC The invention also relates to vaccines including N-terminal Cbpa
CC truncates, which provide protection or elicit protective antibodies
CC to bacterial infection, specifically pneumococcus, and to use of
CC antibodies and antagonists against such polypeptides in diagnosis
CC and passive immunotherapy. The polypeptides and/or polynucleotides
CC are also useful as competitive inhibitors of bacterial adhesion of
CC pneumococcus.

SQ Sequence 428 AA;

Query Match 1.5%; Score 10; DB 20; Length 428;

Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
 |||||
DB 113 ssssssdst 122

RESULT 17

AAV92949
ID AAV92949 standard; Protein; 438 AA.

AC AAV92949;

DT 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 6703.
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis.
XX
XX Corynebacterium glutamicum.
OS
XX
XX EPI108790-A2.
PN
XX
XX 20-JUN-2001.
PD
XX
XX 18-DEC-2000; 2000EP-0127688.
PF
XX
XX 16-DEC-1999; 99JP-0377484.
PR
XX 07-APR-2000; 2000JP-0159162.
PR
XX 03-AUG-2000; 2000JP-0280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PA
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
PI
XX WPI: 2001-376931/40.
DR
XX N-PSDB; AAH68168.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
XX Claim 17; SEQ ID NO: 6703; 246bp + Sequence Listing; English.
PS
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
CC
XX
XX Sequence 438 AA;
SQ

Query Match 1.5%; Score 10; DB 22; Length 438;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 69 sssssdsst 78

RESULT 18
AAV49140
ID AAV49140 standard; Protein; 446 AA.
XX
AC AAV49140;
XX
XX 17-JAN-2000 (first entry)
DT
XX
XX Amino acid sequence of choline-binding protein fragment #4.
DE
XX
XX Truncated surface binding protein; alpha helix; choline binding protein;
KM vaccine; invasive bacterial infection; otitis media; sepsis;
KM meningitis; lobar pneumonia infection; antibody; immature immune system;
KW immunocompromised.

XX Streptococcus pneumoniae.
OS
XX
XX Key Location/Qualifiers
FH
XX
XX Misc-difference 171 /note= "Encoded by AAT"
FT
XX
XX Misc-difference 172 /note= "Encoded by TAC"
FT
XX
XX Misc-difference 173 /note= "Encoded by TTA"
FT
XX
XX Misc-difference 174 /note= "Encoded by CAA"
FT
XX
XX WO951266-A2.
PN
XX
XX 14-OCT-1999.
PD
XX
XX 06-APR-1999; 99WO-US07680.
PF
XX
XX 07-APR-1998; 98US-0080878.
PR
XX 15-MAY-1998; 98US-0085743.
PR
XX
XX (MED-) MEDIMUNE INC.
PA
XX
XX Witzemann TM, Koenig S, Johnson LS;
PI
XX
XX WPI: 1999-601465/51.
DR
XX N-PSDB; AAZ31080.
XX
XX New pneumococcal proteins useful as vaccines and for diagnosis of
PT pneumococcal infections -
PT
XX
XX Claim 10; Page 68-69; 98pp; English.
PS
XX
XX AA49137-Y49152 are amino acid sequences that are fragments of choline
CC binding proteins (CBP). The fragments of the protein are the alpha helix
CC forming parts of the CBPs from Streptococcus pneumoniae. The
CC polypeptides do not contain the actual choline binding fragment. The
CC polypeptides and the nucleotide sequences that encode them
CC (AAZ31077-631092) are used in the invention, which relates to polypeptide
CC truncates of a pneumococcal surface binding protein containing the highly
CC conserved immunogenic alpha helical portion and no choline binding
CC portion. The polypeptides are used as immunogens in a bacterial vaccine.
CC The vaccine can be used for preventing (immunising) or treating invasive
CC bacterial (especially pneumococcal) infections, especially otitis media
CC (caused by S.pneumoniae), sepsis, meningitis and lobar pneumonia
CC infections. Antibodies raised against the polypeptide are useful for
CC detection, prevention (passive immunity) and treatment of S. pneumoniae
CC infections. The vaccines are especially useful in immunocompromised
CC patients, those with an immature immune system, or patients with an on
CC going pneumococcal infection. The vaccine avoids unnecessary expense and
CC provides broad protection against a range of pneumococcal serotypes and
CC it produces an improved and enhanced effect in preventing bacterial
CC infections.
CC
XX
XX Sequence 446 AA;
SQ

Query Match 1.5%; Score 10; DB 20; Length 446;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 114 sssssdsst 123

RESULT 19
AAV49143
ID AAV49143 standard; Protein; 446 AA.
XX
AC AAV49143;
XX
XX

DT 17-JAN-2000 (first entry)
XX
XX Amino acid sequence of choline-binding protein fragment #7.
DE
XX Truncated surface binding protein; alpha helix; choline binding protein;
KM vaccine; invasive bacterial infection; otitis media; sepsis;
KM meningitis; lobar pneumonia infection; antibody; immature immune system;
KM immunocompromised.
XX
OS Streptococcus pneumoniae.
XX
XX WO951266-A2.
XX
XX 14-OCT-1999.
PD
XX 06-APR-1999; 99WO-US07680.
PF
XX 07-APR-1998; 98US-0080878.
PR
XX 15-MAY-1998; 98US-0085743.
PR
XX (MEDI-) MEDIMUNE INC.
PA
XX Wismann TM, Koenig S, Johnson LS;
PI WPI; 1999-601465/51.
XX
XX N-PSDB; AA231083.
DR
XX New pneumococcal proteins useful as vaccines and for diagnosis of
PT pneumococcal infections -
XX
XX Claim 10; Page 72-73; 98pp; English.
PS
XX AA49137-149152 are amino acid sequences that are fragments of choline
CC binding proteins (CBP). The fragments of the protein are the alpha helix
CC forming parts of the CBPs from Streptococcus pneumoniae. The
CC polypeptides do not contain the actual choline binding fragment. The
CC polypeptides and the nucleotide sequences that encode them
CC (AA231077-231092) are used in the invention, which relates to polypeptide
CC truncates of a pneumococcal surface binding protein containing the highly
CC conserved immunogenic alpha helical portion and no choline binding
CC portion. The polypeptides are used as immunogens in a bacterial vaccine.
CC The vaccine can be used for preventing (immunising) or treating invasive
CC bacterial (especially pneumococcal) infections, especially otitis media
CC (caused by S.pneumoniae), sepsis, meningitis and lobar pneumonia
CC infections. Antibodies raised against the polypeptide are useful for
CC detection, prevention (passive immunity) and treatment of S. pneumoniae
CC infections. The vaccines are especially useful in immunocompromised
CC patients, those with an immature immune system, or patients with an on
CC going pneumococcal infection. The vaccine avoids unnecessary expense and
CC provides broad protection against a range of pneumococcal serotypes and
CC it produces an improved and enhanced effect in preventing bacterial
CC infections.
CC
XX
XX Sequence 446 AA;
SQ

Query Match 1.5%; Score 10; DB 20; Length 446;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
DB 114 sssssdst 123

RESULT 20
ID AA49251
AA49251 standard; Protein; 459 AA.
XX
XX AA49251;
XX
XX 07-FEB-2000 (first entry)
DT
XX

DE N-terminal region of CbpA polypeptide ATCC4.
XX
XX Choline binding protein A; CbpA; truncate; immune response; infection;
KM pneumococcal bacterium; vaccine.
KM
XX
OS Streptococcus pneumoniae.
XX
XX WO951187-A2.
XX
XX 14-OCT-1999.
PD
XX 07-APR-1999; 99WO-US07668.
PF
XX 07-APR-1998; 98US-0056019.
PR
XX 07-APR-1998; 98US-0080878.
PR
XX (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
PA
XX Tuomanen EI, Masure HR;
PI WPI; 1999-620161/53.
XX
XX Novel polypeptides, used to develop products for the diagnosis,
PT prevention and treatment of pneumococcal infections -
PT
XX
XX Disclosure; Fig 2A-B; 85pp; English.
PS
XX The invention provides novel isolated polypeptides comprising the amino
CC acid sequence of an N-terminal choline binding protein A (CbpA) truncate
CC (CbpA7). The polypeptides can be selected from sequences shown in
CC AA49225, AA49227, AA49230, AA49231, AA49233 and AA49238. The
CC polypeptides can be used for inducing an immune response in a subject
CC which has been exposed to or infected with a pneumococcal bacterium. They
CC can also be used for preventing infection by a pneumococcal bacterium.
CC Vaccines comprising the polypeptides or encoding nucleic acids can be
CC used for treating a subject infected with or exposed to a pneumococcal
CC bacterium. Antibodies specifically binding the polypeptides can be used
CC for detection and diagnosis and for preventing pneumococcal attachment to
CC a mucosal surface. The products can be used in humans and other animals
CC such as domestic animals, such as feline or canine subjects, farm animals
CC such as bovine, equine, caprine, ovine, and porcine subjects, wild
CC animals (whether in the wild or in a zoological garden), research
CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
CC i.e. for veterinary medical use.
XX
XX Sequence 459 AA;
SQ

Query Match 1.5%; Score 10; DB 20; Length 459;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
DB 127 sssssdst 136

RESULT 21
ID AA492190
AA492190 standard; Protein; 459 AA.
XX
XX AA492190;
XX
XX 01-FEB-2000 (first entry)
DT
XX
XX N-terminal choline binding protein A (CbpA) truncate.
XX
XX Choline binding protein; CbpA; truncate; adhesin; immunogen;
KM vaccine; diagnosis; therapy; pneumococcus; sepsis; meningitis;
KM otitis media; pneumonia.
XX
XX Streptococcus pneumoniae.
OS
OS Synthetic.

XX WO9951188-A2.
 PN
 XX
 PD 14-OCT-1999.
 PT
 XX
 PF 07-APR-1999: 99WO-US07669.
 PR
 XX 07-APR-1998: 98US-0056019.
 PR 07-APR-1998: 98US-0080878.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA (MEDI-) MEDIMUNE INC.
 XX
 PI Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;
 DR WPI, 1999-633690/54.
 XX
 PT New N-terminal choline binding protein A truncate polypeptides, used to
 PT develop products for the diagnosis, prevention and treatment of
 PT pneumococcal infections -
 PS
 XX Claim 47; Fig 2A-B; 160pp; English.
 CC The present sequence represents an N-terminal choline binding
 CC protein A (Cbpa) truncate, denoted ATCC4 Cbpa trun, derived from
 CC Streptococcus pneumoniae Cbpa. Claimed vaccines contain and
 CC N-terminal Cbpa truncate such as the present sequence, or a
 CC polypeptide comprising a conserved region of the Cbpa truncate.
 CC The vaccines provide protection or elicit protective antibodies to
 CC bacterial infection, specifically pneumococcus. Antibodies and
 CC and antagonists against the N-terminal Cbpa truncates are used in
 CC diagnosis and passive immunotherapy.
 XX
 SQ Sequence 459 AA;

Query Match 1.5%; Score 10; DB 20; Length 459;
 Best Local Similarity 100.0%; Pred. No. 0.57;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 616 SSSSSSDSST 625
 |||||||||
 DB 127 ssssssdst 136

RESULT 22
 AAY49250
 ID AAY49250 standard; Protein; 460 AA.
 XX
 AC AAY49250;
 XX
 DT 07-FEB-2000 (first entry)
 XX
 DE N-terminal region of Cbpa polypeptide Ntype4.
 XX
 KW Choline binding protein A; Cbpa; truncate; immune response; infection;
 KW pneumococcal bacterium; vaccine.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO9951187-A2.
 PD 14-OCT-1999.
 XX
 PF 07-APR-1999: 99WO-US07668.
 PR
 XX 07-APR-1998: 98US-0056019.
 PR 07-APR-1998: 98US-0080878.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA
 PI Tuomanen EI, Masure HR;
 PT

DR WPI, 1999-620161/53.
 XX
 PT Novel polypeptides, used to develop products for the diagnosis,
 PT prevention and treatment of pneumococcal infections -
 XX
 PS Disclosure; Fig 2A-B; 85pp; English.
 XX
 CC The invention provides novel isolated polypeptides comprising the amino
 CC acid sequence of an N-terminal choline binding protein A (Cbpa) truncate
 CC (CbpaT). The polypeptides can be selected from sequences shown in
 CC AAY49225, AAY49227, AAY49230, AAY49231, AAY49238, The
 CC polypeptides can be used for inducing an immune response in a subject
 CC which has been exposed to or infected with a pneumococcal bacterium. They
 CC can also be used for preventing infection by a pneumococcal bacterium.
 CC Vaccines comprising the polypeptides or encoding nucleic acids can be
 CC used for treating a subject infected with or exposed to a pneumococcal
 CC bacterium. Antibodies specifically binding the polypeptides can be used
 CC for detection and diagnosis and for preventing pneumococcal attachment to
 CC a mucosal surface. The products can be used in humans and other animals
 CC such as bovine, equine, canine, ovine, and porcine subjects, farm animals
 CC animals (whether in the wild or in a zoological garden), research
 CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
 CC i.e. for veterinary medical use.
 XX
 SQ Sequence 460 AA;

Query Match 1.5%; Score 10; DB 20; Length 460;
 Best Local Similarity 100.0%; Pred. No. 0.57;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 616 SSSSSSDSST 625
 |||||||||
 DB 128 ssssssdst 137

RESULT 23
 AAY32189
 ID AAY32189 standard; Protein; 460 AA.
 XX
 AC AAY32189;
 XX
 DT 01-FEB-2000 (first entry)
 XX
 DE N-terminal choline binding protein A (Cbpa) truncate.
 XX
 KW Choline binding protein; Cbpa; truncate; adhesin; immunogen;
 KW vaccine; diagnosis; therapy; pneumococcus; sepsis; meningitis;
 KW otitis media; pneumonia.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN Synthetic.
 XX
 PD WO9951188-A2.
 XX
 PF 07-APR-1999: 99WO-US07669.
 PR
 XX 07-APR-1998: 98US-0056019.
 PR 07-APR-1998: 98US-0080878.
 XX
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA (MEDI-) MEDIMUNE INC.
 XX
 PI Tuomanen EI, Masure HR, Wizemann TM, Johnson LS, Koenig S;
 DR WPI, 1999-633690/54.
 XX
 PT New N-terminal choline binding protein A truncate polypeptides, used to
 PT develop products for the diagnosis, prevention and treatment of
 PT pneumococcal infections -

XX Claim 47; Fig 2A-B; 160pp; English.
 PS
 CC The present sequence represents an N-terminal choline binding
 CC protein A (Cbpa) truncate, denoted Ntype 4 Cbpa trin, derived from
 CC Streptococcus pneumoniae Cbpa. Claimed vaccines contain and
 CC N-terminal Cbpa truncate such as the present sequence, or a
 CC polypeptide comprising a conserved region of the Cbpa truncate.
 CC The vaccines provide protection or elicit protective antibodies to
 CC bacterial infection, specifically pneumococcus. Antibodies and
 CC and antagonists against the N-terminal Cbpa truncates are used in
 CC diagnosis and passive immunotherapy.
 SQ Sequence 460 AA:
 OY 616 SSSSSSDSST 625
 DB 128 sssssdsst 137
 Query Match 1.5%; Score 10; DB 20; Length 460;
 Best Local Similarity 100.0%; Pred. No. 0.57;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 24
 AAU03646
 ID AAU03646 standard; Protein; 462 AA.
 AC AAU03646;
 DT 12-SEP-2001 (first entry)
 DE Group B Streptococcus antigenic protein, ID-122.
 XX
 KW Group B Streptococcus; encapsulated bacterium; therapeutic; sepsis;
 KW meningitis; neonate; antigenic; vaccine; infection; genital tract;
 KW capsid polysaccharide vaccination.
 OS Streptococcus agalactiae.
 PN WO200132882-A2.
 PD 10-MAY-2001.
 PE 07-SEP-2000; 2000WO-GB03437.
 PR 07-SEP-1999; 99GB-0021125.
 PA (MICR-) MICROBIAL TECHNICS LTD.
 PI Le Page RWF, Wells JM, Hanniffy SB;
 DR WPI: 2001-316444/33.
 DR N-PSDB; AAS07063.
 PT New polypeptides derived from Streptococcus agalactiae are useful to
 PT provide detection of, and vaccination against, Group B Streptococcus
 PT infections, particularly to prevent infection in neonates
 PS Claim 1; Fig 1; 178pp; English.
 XX AAU03601-AAU03722 represent Group B Streptococcus (Streptococcus
 CC agalactiae) amino acid sequences of the invention. S. agalactiae is an
 CC encapsulated bacterium which is a major pathogen of humans causing sepsis
 CC and meningitis in neonates as well as adults. The S. agalactiae antigenic
 CC polypeptides are used to vaccinate against Group B Streptococcus
 CC infections, particularly to prevent infection in new born children
 CC arising from the maternal genital tract. An immunogenic composition is
 CC useful in the preparation of a medicament for the treatment or
 CC prophylaxis of Group B Streptococcus infection. The invention does not
 CC have the disadvantages of varied response rate associated with prior art
 CC capsid polysaccharide vaccination against Group B Streptococcus.

XX sequence 462 AA:
 SQ
 OY 313 RDWGSFMKPI 322
 DB 369 rdwgsfmkpi 378
 Query Match 1.5%; Score 10; DB 22; Length 462;
 Best Local Similarity 100.0%; Pred. No. 0.58;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 25
 AAY49226
 ID AAY49226 standard; Protein; 655 AA.
 AC AAY49226;
 DT 07-FEB-2000 (first entry)
 DE Cbpa of serotype 4 amino acid sequence.
 XX
 KW Choline binding protein A; Cbpa; truncate; immune response; infection;
 KW pneumococcal bacterium; vaccine.
 OS Streptococcus pneumoniae.
 PN WO951187-A2.
 PD 14-OCT-1999.
 PE 07-APR-1999; 99WO-US07668.
 PR 07-APR-1998; 98US-0056019.
 PR 07-APR-1998; 98US-0080878.
 PA (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PI Tuomanen EI, Masure HR;
 DR WPI: 1999-620161/53.
 DR N-PSDB; AAZ31402.
 PT Novel polypeptides, used to develop products for the diagnosis,
 PT prevention and treatment of pneumococcal infections
 PS Disclosure; Page 9; 85pp; English.
 XX The invention provides novel isolated polypeptides comprising the amino
 CC acid sequence of an N-terminal choline binding protein A (Cbpa) truncate
 CC (CbpaT). The polypeptides can be selected from sequences shown in
 CC AAY49225, AAY49227, AAY49230, AAY49231, AAY49233 and AAY49238. The
 CC polypeptides can be used for inducing an immune response in a subject
 CC which has been exposed to or infected with a pneumococcal bacterium. They
 CC can also be used for preventing infection by a pneumococcal bacterium.
 CC Vaccines comprising the polypeptides or encoding nucleic acids can be
 CC used for treating a subject infected with or exposed to a pneumococcal
 CC bacterium. Antibodies specifically binding the polypeptides can be used
 CC for detection and diagnosis and for preventing pneumococcal attachment to
 CC a mucosal surface. The products can be used in humans and other animals
 CC such as domestic animals, such as feline or canine subjects, farm animals
 CC such as bovine, equine, caprine, ovine, and porcine subjects, wild
 CC animals (whether in the wild or in a zoological garden), research
 CC animals, such as mice, rats, rabbits, goats, sheep, pigs, dogs, cats,
 CC i.e. for veterinary medical use.
 SQ Sequence 655 AA:
 OY
 DB
 Query Match 1.5%; Score 10; DB 20; Length 655;
 Best Local Similarity 100.0%; Pred. No. 0.8;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
 1111111111
 DB 113 sssssdsst 122

RESULT 26

AAV32099
 ID AAV32099 standard; Protein; 655 AA.

AC AAV32099;

DT 01-FEB-2000 (first entry)

DE Choline binding protein A (Cbpa).

KM Choline binding protein; Cbpa; adhesin; immunogen;

KW vaccine; diagnosis; therapy; pneumococcus; sepsis; meningitis;

OS otitis media; pneumonia.

XX Streptococcus pneumoniae.

PN WO951188-A2.

PD 14-OCT-1999.

PF 07-APR-1999; 99WO-US07669.

PR 07-APR-1998; 98US-0056019.

PR 07-APR-1998; 98US-0080878.

XX (SJD-) ST JUDE CHILDREN'S RES HOSPITAL.

PA (MEDI-) MEDIMUNE INC.

XX Tuomanen ET, Masure HR, Witzmann TM, Johnson LS, Koenig S;

XX WPI; 1999-633690/54.

DR N-PSDB; AA234452.

XX New N-terminal choline binding protein A truncate polypeptides, used to

PT develop products for the diagnosis, prevention and treatment of

PT pneumococcal infections

XX Disclosure; Page 88-93; 160pp; English.

PS This sequence represents the choline binding protein A (Cbpa)

CC of Streptococcus pneumoniae serotype type 4. The invention provides

CC novel N-terminal Cbpa truncated polypeptides (see AAV32098-110 and

CC AAV32179-90) and polynucleotides encoding them, host-vector systems,

CC and antibodies that specifically bind to the polypeptides. The

CC invention also relates to vaccines including the polypeptides. The

CC which provide protection or elicit protective antibodies to

CC bacterial infection, specifically pneumococcus, and to antibodies

CC and antagonists against such polypeptides for use in diagnosis and

CC passive immunotherapy. The polypeptides and/or polynucleotides are

CC also useful as competitive inhibitors of bacterial adhesion of

CC pneumococcus.

CC Sequence 655 AA:

Query Match 1.5%; Score 10; DB 20; Length 655;

Best Local Similarity 100.0%; Pred. No. 0.8;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625

1111111111

DB 113 sssssdsst 122

RESULT 27
 AAV81653
 ID AAV81653 standard; Protein; 694 AA.

XX AAV81653;
 AC 24-MAY-2000 (first entry)
 DT
 XX

DE Streptococcus pneumoniae protein sequence ID302.

KM Streptococcus pneumoniae; vaccine; screening; protein antigen;

KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;

XX pneumococcal disease.

OS Streptococcus pneumoniae.

PN WO200006737-A2.

PD 10-FEB-2000.

PF 27-JUL-1999; 99WO-GB02451.

PR 27-JUL-1998; 98GB-0016337.

PR 19-MAR-1999; 99US-0125164.

XX (MICR-) MICROBIAL TECHNIQS LTD.

XX Gilbert CFG, Hansbro PM;

XX WPI; 2000-195300/17.

DR New Streptococcal protein, useful as a vaccine, for diagnosis of

PT pneumococcal diseases and for screening agents capable of antagonizing

PT or inhibiting expression of the protein

XX Claim 2; Page 96; 108pp; English.

PS AAV81501 to AAV81679 represent specifically claimed protein sequences

CC isolated from Streptococcus pneumoniae. AAV05407 to AAV05590 represent

CC specifically claimed nucleotide sequences isolated from S. pneumoniae.

CC The sequences have antibacterial and antiinflammatory properties.

CC The protein sequences, and fragments of them, are useful as immunogens

CC and/or antigens. The nucleotide sequences can be used in vaccines and in

CC diagnostic assays. The proteins and nucleotides can be useful for the

CC detection and diagnosis of S. pneumoniae. The protein sequences are also

CC useful for screening an agent capable of antagonizing, inhibiting or

CC interfering with the function or expression of the proteins in which the

CC agent is useful for treatment or prophylaxis of S. pneumoniae infection

CC and meningitis. AAV05591 to AAV05614 represent primers used in the

CC exemplification of the present invention.

XX Sequence 694 AA:

Query Match 1.5%; Score 10; DB 21; Length 694;

Best Local Similarity 100.0%; Pred. No. 0.84;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625

1111111111

DB 151 sssssdsst 160

RESULT 28

AAU36453
 ID AAU36453 standard; Protein; 774 AA.

AC AAU36453;

DT 14-FEB-2002 (first entry)

DE Pseudomonas aeruginosa cellular proliferation protein #443.

XX Antisense; prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

XX


```

XX OS Heliothis armigera stunt virus.
XX PN WO9404660-A.
XX PD 03-MAR-1994.
XX PF 13-AUG-1993; 93WO-AU00411.
XX PR 14-AUG-1992; 92AU-0004081.
XX PR 08-JUL-1993; 93US-0089372.
XX PA (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX PA (PACI-) PACIFIC SEEDS PTY LTD.
XX PI Christian PD, Gordon KHJ, Hanzlik TN;
XX DR WPI, 1994-083180/10.
XX DR N-PSDB; AAQ58522.
XX PT Small RNA virus capable of infecting insect species, e.g.
XX PT Heliothis - and transgenic plants contg. viral nucleic acid, for
XX PT protection against insect pests
XX PS Disclosure; Figure 1, 183pp; English.
XX CC The inventors claim a virus comprising a genome hybridisable with
XX CC the nucleotide sequence of RNA 1 or RNA 2; pref. the sequences are
XX CC those given in Figs 1 and 2 of the specification. Isolated proteins
XX CC or polypeptide prepn. of the proteins or polypeptides derivable
XX CC from the virus are also claimed.
XX CC H. armigera larvae were raised and viral RNA was extracted. The virus
XX CC RNAs were reverse transcribed into cDNA. Clone E3 represents 99.7%
XX CC of RNA 1.(hr236 contains about 888 or RNA 2.)The full length clone
XX CC of RNA 1 was completed using PCR. RNA 1 encodes a protein of mol.
XX CC wt. 187,000 which is regarded as the replicase in view of its AA
XX CC sequence similarity in certain limited regions to replicases of
XX CC of other RNA viruses. The apparent mol. wt. of this protein upon
XX CC in vitro translation of virus RNA and SDS-PAGE is 195,000. The
XX CC sequence given in Figure 1 is in the same sense as the viral
XX CC (positive-sense) RNA. There are other small ORFs at the 3' end,
XX CC corresp. to the proteins P1a, P1b and P14.
XX SQ Sequence 1704 AA;

```

```

Query Match 1.5%; Score 10; DB 15; Length 1704;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 617 SSSSDSSTS 626
   |||||||||
DB 1452 ssssdsts 1461

```

Search completed: June 13, 2002, 08:46:22
 Job time: 263 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:45:24 ; Search time 21.51 Seconds
(without alignments)
2975.151 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666

Sequence: 1 KTYDNKNQLIADLGSERVRN.....TQGSNTTPDQGNQNPQAPQ 666

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 10

Total number of hits satisfying chosen parameters: 33

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	666	100.0	719	2	G95042
2	502	75.4	719	2	S28036
3	351	52.7	608	2	S28036
4	332	49.8	719	2	S28037
5	330	49.5	719	2	A42893
6	330	49.5	719	2	A97913
7	181	27.2	719	2	S28031
8	181	27.2	719	2	S28034
9	181	27.2	719	2	S28033
10	181	27.2	719	2	S28032
11	97	14.6	188	2	S31952
12	79	11.9	719	2	S28035
13	53	8.0	637	2	B42893
14	20	3.0	139	2	S31941
15	14	2.1	664	2	G66692
16	12	1.8	760	2	E84953
17	11	1.7	801	2	B86673
18	11	1.7	821	2	E95245
19	11	1.7	821	2	B99110
20	11	1.7	966	2	E84053
21	10	1.5	504	2	T29028
22	10	1.5	504	2	T21377
23	10	1.5	577	1	A36442
24	10	1.5	693	2	H95255
25	10	1.5	730	2	A13460
26	10	1.5	757	2	G97472
27	10	1.5	757	2	AC2691
28	10	1.5	764	2	B97371
29	10	1.5	764	2	AB2589

30	10	1.5	774	2	H83057	penicillin-binding
31	10	1.5	777	2	A82303	penicillin-binding
32	10	1.5	781	2	D64138	penicillin-binding
33	10	1.5	824	2	AC0412	penicillin-binding

ALIGNMENTS

RESULT 1

G95042
penicillin-binding protein 1A [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
A:Accession: G95042
R:Retcelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: G95042
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK74536.1; PID:q14971838; GSPDB:GN00164; TIGR:
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0369
C:Superfamily: penicillin-binding protein 1B

Query Match 100.0%; Score 666; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KTYDNKNQLIADLGSERVRNAQNDIPDLVKAIVSIEDHREDFHRCIDTITILGAFUN	60
DB	54	KTYDNKNQLIADLGSERVRNAQNDIPDLVKAIVSIEDHREDFHRCIDTITILGAFUN	113
QY	61	LOSNSLGGSTLTQOLIKLTFSTSTSDQTTISRRQEAWLAIQDQKATKQELIYYINK	120
DB	114	LOSNSLGGSTLTQOLIKLTFSTSTSDQTTISRRQEAWLAIQDQKATKQELIYYINK	173
QY	121	YMSGNGTGMOTAQNYGKDLNLSLPQALLAGMPQAPNOYDPSHPEAAQDRRLVL	180
DB	174	YMSGNGTGMOTAQNYGKDLNLSLPQALLAGMPQAPNOYDPSHPEAAQDRRLVL	233
QY	181	SMKMGYISAOYERKAVNPTTDDGLQSLKSASNPAYMDNLTKEVINVEBETGYNLT	240
DB	234	SMKMGYISAOYERKAVNPTTDDGLQSLKSASNPAYMDNLTKEVINVEBETGYNLT	293
QY	241	TGMDVYTNVDDQAKHLMIDYNTDEVVAYPPDELOVASTIVVNSGKVAQAQARHOSN	300
DB	294	TGMDVYTNVDDQAKHLMIDYNTDEVVAYPPDELOVASTIVVNSGKVAQAQARHOSN	353
QY	301	VSEGINQAVETNRDGSSTMKPTTDYAPALEYGVYDSTATIVHDEBYNPGTNPVYNNDR	360
DB	354	VSEGINQAVETNRDGSSTMKPTTDYAPALEYGVYDSTATIVHDEBYNPGTNPVYNNDR	413
QY	361	GYFGNITLQYALQOSRNPAYETLKKVGINRAKTLNLGLIDYPSIHYSNATISSTTESD	420
DB	414	GYFGNITLQYALQOSRNPAYETLKKVGINRAKTLNLGLIDYPSIHYSNATISSTTESD	473
QY	421	KKYGSSSEKMAAAYAFANGGYTPMYTHKVPFSDGSEKESNNGTAMKETTAYMMTD	480
DB	474	KKYGSSSEKMAAAYAFANGGYTPMYTHKVPFSDGSEKESNNGTAMKETTAYMMTD	533
QY	481	MMKTVLTGTRNAYLAWLPQAGKTGTSNYTDEELENIKTSQFVAPDELPAGYTRKYSM	540
DB	534	MMKTVLTGTRNAYLAWLPQAGKTGTSNYTDEELENIKTSQFVAPDELPAGYTRKYSM	593

OY 541 AWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFRNGAR 600
|||||
Db 594 AWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFRNGAR 653
|||||
OY 601 STWNSAPPOOPSTESSSSSDSTSSSTPTSTNNSTTTNNNTQOQNTTPOOQON 660
|||||
Db 654 STWNSAPPOOPSTESSSSSDSTSSSTPTSTNNSTTTNNNTQOQNTTPOOQON 713
|||||
OY 661 POPAP 666
|||||
Db 714 POPAP 719

RESULT 2

S28038
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 45607) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28038
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of pen
A:Reference number: S28031; MIDID:93010977
A:Status: translation not shown
A:Accession: S28038
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67873; NID:g47419; PIDN:CAA48073.1; PID:g47420
C:Superfamily: penicillin-binding protein 1B

Query Match 75.4%; Score 502; DB 2; Length 719;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 602; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFARN 60
|||||
Db 54 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFARN 113
|||||
OY 61 LOSNSLOGGSTLTQOLIKLTFSTSDOTISRAQEAWLAIOLBKATKQELITYYINK 120
|||||
Db 114 LOSNSLOGGSTLTQOLIKLTFSTSDOTISRAQEAWLAIOLBKATKQELITYYINK 173
|||||
OY 121 VYMSNGYNGQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRNLVL 180
|||||
Db 174 VYMSNGYNGQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRNLVL 253
|||||
OY 181 SEKNNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEYINOVEEETGYNLLT 240
|||||
Db 234 SEKNNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEYINOVEEETGYNLLT 293
|||||
OY 241 TGMDEVYTNVDOEAQKHMDIYNTDEVYVAYPPDELOVASTIVDSNGKVIAQLGARHOSSN 300
|||||
Db 294 TGMDEVYTNVDOEAQKHMDIYNTDEVYVAYPPDELOVASTIVDSNGKVIAQLGARHOSSN 353
|||||
OY 301 VSFGINQAVETNRDWSSTMKPITDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNMDR 360
|||||
Db 354 VSFGINQAVETNRDWSSTMKPITDYAPALEYGYDSTATIVHDEPNYPGTNTPVYNMDR 413
|||||
OY 361 GFYGNITLQYALQOQRNVAVERLNVGNLRATFLNGLIDYPSIHYSAISSNTTESD 420
|||||
Db 414 GFYGNITLQYALQOQRNVAVERLNVGNLRATFLNGLIDYPSIHYSAISSNTTESD 473
|||||
OY 421 KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYVSDGSEKEFSNVGTRAKKETATAMTE 480
|||||
Db 474 KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYVSDGSEKEFSNVGTRAKKETATAMTE 533
|||||
OY 481 MKKTVLTGTGRNAYLAMPQAGKTGTSVYTDIEINHLKTSOFVAPDELFAGYTRKYSM 540
|||||
Db 534 MKKTVLTGTGRNAYLAMPQAGKTGTSVYTDIEINHLKTSOFVAPDELFAGYTRKYSM 593
|||||
OY 541 AWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFRNGAR 600
|||||

Db 594 AWTGYSNRLTPLVNGLTVAARVYRSMYTLSEGSNPEDMNIPEGLYRNGEEVFRNGAR 653
OY 601 STW 603
Db 654 STW 656

RESULT 3

S28036
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 8250) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28036
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MIDID:93010977
A:Accession: S28036
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-608 <MAR>
A:Cross-references: EMBL:X67871; NID:g47415; PIDN:CAA48071.1; PID:g47416
C:Superfamily: penicillin-binding protein 1B

Query Match 52.7%; Score 351; DB 2; Length 608;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFARN 60
|||||
Db 54 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDITRILGAFARN 113
|||||
OY 61 LOSNSLOGGSTLTQOLIKLTFSTSDOTISRAQEAWLAIOLBKATKQELITYYINK 120
|||||
Db 114 LOSNSLOGGSTLTQOLIKLTFSTSDOTISRAQEAWLAIOLBKATKQELITYYINK 173
|||||
OY 121 VYMSNGYNGQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRNLVL 180
|||||
Db 174 VYMSNGYNGQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAADRNLVL 233
|||||
OY 181 SEKNNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEYINOVEEETGYNLLT 240
|||||
Db 234 SEKNNGYISAEQYERAVNPTITDGLQSLKSASNPAYMDNYLKEYINOVEEETGYNLLT 293
|||||
OY 241 TGMDEVYTNVDOEAQKHMDIYNTDEVYVAYPPDELOVASTIVDSNGKVIAQLGARHOSSN 300
|||||
Db 294 TGMDEVYTNVDOEAQKHMDIYNTDEVYVAYPPDELOVASTIVDSNGKVIAQLGARHOSSN 353
|||||
OY 301 VSFGINQAVETNRDWSSTMKPITDYAPALEYGYDSTATIVHDEPNYPGT 351
|||||
Db 354 VSFGINQAVETNRDWSSTMKPITDYAPALEYGYDSTATIVHDEPNYPGT 404
|||||

RESULT 4

S28037
penicillin-binding protein 1a - Streptococcus pneumoniae (strain 63915) (fragment)
C:Species: Streptococcus pneumoniae
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28037
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of
A:Reference number: S28031; MIDID:93010977
A:Accession: S28037
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67872; NID:g47417; PIDN:CAA48072.1; PID:g47418
C:Superfamily: penicillin-binding protein 1B

Query Match 49.8%; Score 332; DB 2; Length 719;

Best Local Similarity 100.0%; Pred. No. 0; Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 KIYDKNKLIDLCSEKRVNAQANDIPDLVKAIVSIEDHFFDRGIDITRILGAFLRN 60
Db 54 KIYDNKNLIDLCSERVRVNAQANDIPDLVKAIVSIEDHFFDRGIDITRILGAFLRN 113
OY 61 LOSNLSAGSFLTQOOLIKLYFSTSTSDQTSRKAQEWMLAIOLEOKATKOEILTYINK 120
Db 114 LOSNLSAGSFLTQOOLIKLYFSTSTSDQTSRKAQEWMLAIOLEOKATKOEILTYINK 173
OY 121 VYMSNGNMGTAQNTAONYGKDLNLSLPQALLAGMPAPNOYDPSHPEAAQDRNLVL 180
Db 174 VYMSNGNMGTAQNTAONYGKDLNLSLPQALLAGMPAPNOYDPSHPEAAQDRNLVL 233
OY 181 SEMKNOGYTSAEYKAVNPTITDGLQSLKASNPAYMDNLYLKEVINOVEEETGNLLT 240
Db 234 SEMKNOGYTSAEYKAVNPTITDGLQSLKASNPAYMDNLYLKEVINOVEEETGNLLT 293
OY 241 TGMVYTNVDOEAQKHLMDIYNTDEYVAPDDELQVASTIVDSNGKVIQAOLGARHOSN 300
Db 294 TGMVYTNVDOEAQKHLMDIYNTDEYVAPDDELQVASTIVDSNGKVIQAOLGARHOSN 353
OY 301 VSPGINQAVETNRDWSGTMRPTTDYAPALEYG 332
Db 354 VSPGINQAVETNRDWSGTMRPTTDYAPALEYG 385

RESULT 5
A42893
penicillin-binding protein 1A - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Jan-2000
C:Accession: A42893
R:Martin, C.; Briese, T.; Hakenbeck, R.
J. Bacteriol. 174, 4517-4523, 1992
A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae
A:Reference number: A42893; MUID:92325042
A:Accession: A42893
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: GB:M90527; NID:g153766; PIDN:AAA26956.1; PID:g153768
C:Superfamily: penicillin-binding protein 1B
```

Query Match 49.5%; Score 330; DB 2; Length 719; Best Local Similarity 99.6%; Pred. No. 0; Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
OY 72 LTQOLIKLYFSTSTSDQTSRKAQEWMLAIOLEOKATKOEILTYINKVYMSNGNYGMQ 131
Db 125 LTQOLIKLYFSTSTSDQTSRKAQEWMLAIOLEOKATKOEILTYINKVYMSNGNYGMQ 184
OY 132 TAAONYGKDLNLSLPQALLAGMPAPNOYDPSHPEAAQDRNLVLSEKKNQGYTSA 191
Db 185 TAAONYGKDLNLSLPQALLAGMPAPNOYDPSHPEAAQDRNLVLSEKKNQGYTSA 244
OY 192 EYKAVNPTITDGLQSLKASNPAYMDNLYLKEVINOVEEETGNLLTGMVYTNVDO 251
Db 245 EYKAVNPTITDGLQSLKASNPAYMDNLYLKEVINOVEEETGNLLTGMVYTNVDO 304
OY 252 EAKHLMIDYNTDEYVAPDDELQVASTIVDSNGKVIQAOLGARHOSNVSFGINQAVET 311
Db 305 EAKHLMIDYNTDEYVAPDDELQVASTIVDSNGKVIQAOLGARHOSNVSFGINQAVET 364
OY 312 NRDMGSTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNPYVMDRGYFGNITLQYA 371
Db 365 NRDMGSTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNPYVMDRGYFGNITLQYA 424
OY 372 LQOSRNVAVETLNKVGILNRKATFLNGIGIDYPSIHYSNAISSNTESDKKYGASSEKMA 431
```

```
Db 425 LQOSRNVAVETLNKVGILNRKATFLNGIGIDYPSIHYSNAISSNTESDKKYGASSEKMA 484
OY 432 AAYAFANGGTYRPMYTHKVFSDGSEKFSNVGTRAMKETTAAYMTDMKTVLYTGTG 491
Db 485 AAYAFANGGTYRPMYTHKVFSDGSEKFSNVGTRAMKETTAAYMTDMKTVLYTGTG 544
OY 492 RNAYLAWLPQAGKTGTSNVTDEIEHNHKTQSFVAPDELFGAGYTRKYSMAVWTGYSNRLT 551
Db 545 RNAYLAWLPQAGKTGTSNVTDEIEHNHKTQSFVAPDELFGAGYTRKYSMAVWTGYSNRLT 604
OY 552 PLVNGGLTVAAKVYRSMNTYLSGSPEDMNPBELYNGEYFKNGARSTW 603
Db 605 PLVNGGLTVAAKVYRSMNTYLSGSPEDMNPBELYNGEYFKNGARSTW 656
```

```
RESULT 6
A97913
peptidoglycan glycosyltransferase (EC 2.4.1.129) [imported] - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: A97913
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnthen, S.; y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: A97913
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-719 <KUR>
A:Cross-references: GB:A007317; PIDN:AAK9133.1; PID:g15457886; GSPDB:GN00174
A:Genetics:
A:Gene: pbpa
C:Superfamily: penicillin-binding protein 1B
C:Keywords: glycosyltransferase; hexosyltransferase
```

Query Match 49.5%; Score 330; DB 2; Length 719; Best Local Similarity 99.6%; Pred. No. 0; Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
OY 72 LTQOLIKLYFSTSTSDQTSRKAQEWMLAIOLEOKATKOEILTYINKVYMSNGNYGMQ 131
Db 125 LTQOLIKLYFSTSTSDQTSRKAQEWMLAIOLEOKATKOEILTYINKVYMSNGNYGMQ 184
OY 132 TAAONYGKDLNLSLPQALLAGMPAPNOYDPSHPEAAQDRNLVLSEKKNQGYTSA 191
Db 185 TAAONYGKDLNLSLPQALLAGMPAPNOYDPSHPEAAQDRNLVLSEKKNQGYTSA 244
OY 192 EYKAVNPTITDGLQSLKASNPAYMDNLYLKEVINOVEEETGNLLTGMVYTNVDO 251
Db 245 EYKAVNPTITDGLQSLKASNPAYMDNLYLKEVINOVEEETGNLLTGMVYTNVDO 304
OY 252 EAKHLMIDYNTDEYVAPDDELQVASTIVDSNGKVIQAOLGARHOSNVSFGINQAVET 311
Db 305 EAKHLMIDYNTDEYVAPDDELQVASTIVDSNGKVIQAOLGARHOSNVSFGINQAVET 364
OY 312 NRDMGSTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNPYVMDRGYFGNITLQYA 371
Db 365 NRDMGSTMKPTTDYAPALEYGVYDSTATIVHDEPNYPGTNPYVMDRGYFGNITLQYA 424
OY 372 LQOSRNVAVETLNKVGILNRKATFLNGIGIDYPSIHYSNAISSNTESDKKYGASSEKMA 431
Db 425 LQOSRNVAVETLNKVGILNRKATFLNGIGIDYPSIHYSNAISSNTESDKKYGASSEKMA 484
OY 432 AAYAFANGGTYRPMYTHKVFSDGSEKFSNVGTRAMKETTAAYMTDMKTVLYTGTG 491
Db 485 AAYAFANGGTYRPMYTHKVFSDGSEKFSNVGTRAMKETTAAYMTDMKTVLYTGTG 544
OY 492 RNAYLAWLPQAGKTGTSNVTDEIEHNHKTQSFVAPDELFGAGYTRKYSMAVWTGYSNRLT 551
```



```

OY 51 IRIIAFLRNQSNLQSGSTLTQOLIKLTFFSTSTQOTISRKQGEWMLAIOLEQKATK 110
Db 104 IRIIGAFRLNQSNSLQSGSTLTQOLIKLTFFSTSTQOTISRKQGEWMLAIOLEQKATK 163
OY 111 QEILTYITNKRYMSNGNMGOTAAQNYGKDLNNLSLPOLLALAGMPAPAPNOYDPYSHE 170
Db 164 QEILTYITNKRYMSNGNMGOTAAQNYGKDLNNLSLPOLLALAGMPAPAPNOYDPYSHE 223
OY 171 AAQDRNRLVLESMKNQGYISAQVEKAVNPTITDGLQSLKASNPAYMDNYLKEVINQV 230
Db 224 AAQDRNRLVLESMKNQGYISAQVEKAVNPTITDGLQSLKASNPAYMDNYLKEVINQV 283
OY 231 E 231
Db 284 E 284

RESULT 11
S31952
penicillin-binding protein 1A - Streptococcus pneumoniae (strain 85983) (fragment)
C:Species: Streptococcus pneumoniae
A:Variety: strain 85983
C:Date: 06-Jan-1995 #sequence_revision 26-Jul-1996 #text_change 26-May-2000
C:Accession: S31952
R:Kell, C.M.; Jordens, Z.; Daniels, M.; Coffey, T.J.; Bates, J.; Paul, J.; Gilks, C.; Se
Submitted to the EMBL Data Library, February 1993
A:Description: Molecular epidemiology of penicillin-resistant pneumococci isolated in Ne
A:Reference number: S31941
A:Accession: S31952
A:Molecule type: DNA
A:Residues: 1-188 <KEL>
A:Cross-references: EMBL:Z21800
A:Experimental source: strain 85983
C:Superfamily: penicillin-binding protein 1B
C:Keywords: antibiotic resistance; penicillin resistance

Query Match 14.6%; Score 97; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 3,4e-90;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 557 GILTAAKYRRSMNTYLSGSDPEDMNIPEGILYRNGEYFKNGARSTWNSPAPQPPSTES 616
Db 92 GILTAAKYRRSMNTYLSGSDPEDMNIPEGILYRNGEYFKNGARSTWNSPAPQPPSTES 151
OY 617 SSSSDSSTSOSSSTPSTNNSTTNPNNNTQOOSNTT 653
Db 152 SSSSDSSTSOSSSTPSTNNSTTNPNNNTQOOSNTT 188

RESULT 12
S28035
penicillin-binding protein 1A - Streptococcus pneumoniae (strain 2039) (fragment)
C:Species: Streptococcus pneumoniae
A:Variety: strain 2039
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 26-May-2000
C:Accession: S28035
R:Martin, C.; Sibold, C.; Hakenbeck, R.
EMBO J. 11, 3831-3836, 1992
A:Title: Relatedness of penicillin-binding protein 1a genes from different clones of pen
A:Reference number: S28031; MUID:93010977
A:Accession: S28035
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-719 <MAR>
A:Cross-references: EMBL:X67870; NID:g47413; PIDN:CAAA8070.1; PID:g47414
C:Experimental source: strain 2039
C:Superfamily: penicillin-binding protein 1B
C:Keywords: antibiotic resistance; penicillin resistance

Query Match 11.9%; Score 79; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 2,2e-71;

```

```

QY      64 NSLGGSTLTQOLIKLTFEFTSTSDQTISRKAQEMALAIQLEOKATKQELIYYINKVYM 123
      |||||||
Db      117 NSLGGSGSTLTQOLIKLTFEFTSTSDQTISRKAQEMALAIQLEOKATKQELIYYINKVYM 176
      |||||||

QY      124 SNGNTGMOTAAONYGKDL 142
      |||||||
Db      177 SNGNYGMOTAAONYGKDL 195
      |||||||

RESULT 13
Ba2893
penicillin-binding protein 1 - Streptococcus oralis
C:Species: Streptococcus oralis
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 04-Mar-2000
C:Accession: BA2893
R:Martin, C.; Briese, T.; Hakenbeck, R.
J. Bacteriol. 174, 4517-4523, 1992
A:Title: Nucleotide sequences of genes encoding penicillin-binding proteins from Stre
A and 1B.
A:Reference number: A42893; MUID:92325042
A:Accession: BA2893
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-637 <MAR>
A:Cross-references: GB:M90528; NID:q153769; PIDN:AA26958.1; PID:q153771
C:Superfamily: penicillin-binding protein 1B

Query Match      8.0%; Score 53; DB 2; Length 637;
Best Local Similarity 100.0%; Pred. No. 4,9e-45;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      266 GKVINOLGARHSSNVSGINQAVETNRDMGSTMKEPITDYAPALEYGYDSTA 338
      |||||||
Db      340 GKVINOLGARHSSNVSGINQAVETNRDMGSTMKEPITDYAPALEYGYDSTA 392
      |||||||

RESULT 14
S31941
penicillin-binding protein 1A - Streptococcus pneumoniae (strain 100511) (fragment)
C:Species: Streptococcus pneumoniae
A:Variety: strain 100511
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-May-2000
C:Accession: S31941
R:Kell, C.M.; Jordens, Z.; Daniels, M.; Coffey, T.J.; Bates, J.; Paul, J.; Gilks, C.;
submitted to the EMBL Data Library, February 1993
A:Description: Molecular epidemiology of penicillin-resistant pneumococci isolated in
A:Reference number: S31941
A:Accession: S31941
A:Molecule type: DNA
A:Residues: 1-139 <KE>
A:Cross-references: EMBL:Z21809; NID:q49384; PIDN:CAA79874.1; PID:q940871
A:Experimental source: strain 100511
C:Superfamily: penicillin-binding protein 1B
C:Keywords: antibiotic resistance; penicillin resistance

Query Match      3.0%; Score 20; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 4.1e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      533 GYTRKYSMAVMTGYSNRLTP 552
      |||||||
Db      16 GYTRKYSMAVMTGYSNRLTP 35
      |||||||

RESULT 15
G86692
penicillin-binding protein 1A [imported] - Lactococcus lactis subsp. lactis (strain I
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

```

C:Accession: G86692
R:Botlittin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664 <STO>
A:Cross-references: GB:AE005176; PID:g12723428; PIDN:AAK04641.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: pona
C:Superfamily: penicillin-binding protein 1B

Query Match 2.1%; Score 14; DB 2; Length 664;
Best Local Similarity 100.0%; Pred. No. 2,1e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 GASSEKMAAYAA 437
|||||
DB 467 GASSEKMAAYAA 480

RESULT 16
E84953
penicillin-binding protein 1b [imported] - *Buchnera* sp. (strain APS)
C:Species: *Buchnera* sp.
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: E84953
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.
Nature 407, 81-86, 2000
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. A
A:Reference number: A84930; MUID:20445173
A:Accession: E84953
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-760 <STO>
A:Cross-references: GB:AP000398; GSPDB:GN00144
A:Experimental source: strain APS
C:Genetics:
A:Gene: mrcB; BU200

Query Match 1.8%; Score 12; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOLIK 78
|||||
DB 204 QGGSTLTQOLIK 215

RESULT 17
B86673
penicillin-binding protein 1b [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL14
C:Species: *Lactococcus lactis* subsp. *lactis*
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: B86673
R:Botlittin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: B86673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-801 <STO>
A:Cross-references: GB:AE005176; PID:g12723258; PIDN:AAK04484.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: pbp1b

Query Match 1.7%; Score 11; DB 2; Length 801;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
|||||
DB 172 GGSTLTQOLIK 182

RESULT 18
E95245
penicillin-binding protein 1b [imported] - *Streptococcus pneumoniae* (strain TIGR4)
C:Species: *Streptococcus pneumoniae*
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: E95245
R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umeyama, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapp
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: E95245
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-821 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK76158.1; PID:g14973609; GSPDB:GN00164; TIGR
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2099

Query Match 1.7%; Score 11; DB 2; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
|||||
DB 176 GGSTLTQOLIK 186

RESULT 19
B99110
peptidoglycan glycosyltransferase (EC 2.4.1.129) [imported] - *Streptococcus pneumonia*
C:Species: *Streptococcus pneumoniae*
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: B99110
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaaskunas, S.
A:Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: B99110
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-821 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAL00711.1; PID:g15459604; GSPDB:GN00174
A:Genetics:
A:Gene: pbp1b
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 1.7%; Score 11; DB 2; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
|||||
DB 176 GGSTLTQOLIK 186

```
RESULT 20
E84053
penicillin-binding protein 1A pdpf [Imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: E84053
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E84053
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-966 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; NID:g10175792; PIDN:BAB06948.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: pdpf

Query Match 1.7%; Score 11; DB 2; Length 966;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
|||||
DB 157 GGSTLTQOLIK 167

RESULT 21
T29028
hypothetical protein F53612.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29028
R:Mu, X.; Graves, T.
Submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid F53612.
A:Reference number: Z20555
A:Accession: T29028
A:Cross-references: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-294 <WUX>
A:Status: preliminary
A:Experimental source: strain Bristol N2; clone F53612
C:Genetics:
A:Gene: CESP:F53612.4
A:Map position: 1
A:Introns: 111/3; 271/1

Query Match 1.5%; Score 10; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 SSSSSSDSTS 626
|||||
DB 126 SSSSSSDSTS 135

RESULT 22
T21377
hypothetical protein F25H9.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21377
R:Smey, R.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19414
A:Accession: T21377
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

A:Residues: 1-504 <ML>
A:Cross-references: EMBL:Z81069; PIDN:CAB02995.1; GSPDB:GN00023; CESP:F25H9.5
A:Experimental source: clone F25H9
C:Genetics:
A:Gene: CESP:F25H9.5
A:Map position: 5
A:Introns: 7/1; 34/3; 88/1; 189/3; 277/2; 343/3; 442/3
C:Superfamily: Caenorhabditis elegans hypothetical protein F25H9.5

Query Match 1.5%; Score 10; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 617 SSSSSSDSTS 626
|||||
DB 443 SSSSSSDSTS 452

RESULT 23
A36442
mitochondrial processing peptidase (EC 3.4.24.64) alpha chain precursor - Neurospora
N:Alternate names: alpha-MPP; mitochondrial processing peptidase catalytic chain; MPP
C:Species: Neurospora crassa
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 19-Jan-2001
C:Accession: A36442; S36362
R:Schneider, H.; Arretz, M.; Wachter, E.; Neupert, W.
J. Biol. Chem. 265, 9881-9887, 1990
A:Title: Matrix processing peptidase of mitochondria. Structure-function relationship
A:Reference number: A36442; MUID:90277682
A:Accession: A36442
A:Molecule type: mRNA
A:Residues: 1-577 <SCH1>
A:Cross-references: GB:J05484
R:Schneider, H.; Arretz, M.; Wachter, E.; Neupert, W.
Submitted to the EMBL Data Library, July 1990
A:Reference number: S36362
A:Accession: S36362
A:Molecule type: mRNA
A:Residues: 1-106, A, 108-577 <SCH2>
A:Cross-references: GB:J05484; NID:g168840; PIDN:AAA33597.1; PID:g168841
C:Genetics:
A:Gene: MPP
A:Genome: nuclear
C:Superfamily: mitochondrial processing peptidase alpha chain
C:Keywords: heterodimer; hydrolase; metalloproteinase; mitochondrial matrix; mitochon
F:1-35/Domain: transit peptide (mitochondrion) #status predicted <TP>
F:36-577/Product: mitochondrial processing peptidase alpha chain #status predicted <M

Query Match 1.5%; Score 10; DB 1; Length 577;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 615 ESSSSSDSS 624
|||||
DB 286 ESSSSSDSS 295

RESULT 24
H95255
choline binding protein A [Imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: H95255
R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95255
```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-693 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK76241.1; PID:g14973701; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2190

Query Match 1.5%; Score 10; DB 2; Length 693;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
|||||||
DB 151 SSSSSSDSST 160

RESULT 25
A13480
penicillin-binding protein 1A [imported] - *Brucella melitensis* (strain 16M)
C:Species: *Brucella melitensis*
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: A13480
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A:Reference number: AB2252; PMID:11756688
A:Accession: A13480
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-730 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL5012.1; PID:g17963868; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11831
A:Map position: 1

Query Match 1.5%; Score 10; DB 2; Length 730;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOL 76
|||||||
DB 146 QGGSTLTQOL 155

RESULT 26
G97472
penicillin-binding protein 1A [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon
C:Species: *Agrobacterium tumefaciens*
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: G97472
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*
A:Reference number: A97359; PMID:11743194
A:Accession: G97472
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK6736.1; PID:g15155930; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_1696
A:Map position: circular chromosome

Query Match 1.5%; Score 10; DB 2; Length 757;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOL 76
|||||||
DB 193 QGGSTLTQOL 202

RESULT 27
AC2691
penicillin binding protein [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont
C:Species: *Agrobacterium tumefaciens*
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AC2691
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCl
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AC2691
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL41945.1; PID:g17739313; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0931
A:Map position: circular chromosome

Query Match 1.5%; Score 10; DB 2; Length 757;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOL 76
|||||||
DB 193 QGGSTLTQOL 202

RESULT 28
B97371
penicillin-binding protein 1A [imported] - *Agrobacterium tumefaciens* (strain C58, Cer
C:Species: *Agrobacterium tumefaciens*
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: B97371
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*
A:Reference number: A97359; PMID:11743194
A:Accession: B97371
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-764 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK85923.1; PID:g15154974; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_158
A:Map position: circular chromosome

Query Match 1.5%; Score 10; DB 2; Length 764;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 QGGSTLTQOL 76
|||||||
DB 183 QGGSTLTQOL 192

RESULT 29
AB2589
penicillin-binding protein pbpc [imported] - *Agrobacterium tumefaciens* (strain C58, D

C:Species: *Agrobacterium tumefaciens*
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AB2589
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCell
; Kap, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AB2589
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-764 <KUR>
A:Cross-references: GB:AE008688; PIDN:AL41128.1; PID:g17738422; GSPDB:GN00186
C:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: pbpc
A:Map position: circular chromosome

Query Match 1.5%; Score 10; DB 2; Length 764;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 OGGSTRTQOL 76
|||||
Db 183 OGGSTRTQOL 192

RESULT 30
H83057
penicillin-binding protein 1B PA4700 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83057
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337
A:Accession: H83057
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-774 <STO>
A:Cross-references: GB:AE004884; GB:AE004091; NID:g9950954; PIDN:AAG08086.1; GSPDB:GN001
C:Genetics:
A:Experimental source: strain PA01
A:Gene: mrcB; PA4700

Query Match 1.5%; Score 10; DB 2; Length 774;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 OGGSTRTQOL 76
|||||
Db 220 OGGSTRTQOL 229

RESULT 31
AB2303
penicillin-binding protein 1B VC0602 [imported] - *Vibrio cholerae* (strain N16961 serogro
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: AB2303
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, R
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A:Reference number: A82035; MUID:20406833
A:Accession: AB2303
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-777 <HEI>
A:Cross-references: GB:AE004145; GB:AE003852; NID:g9655032; PIDN:AAF93769.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0602
A:Map position: 1
C:Superfamily: penicillin-binding protein 1B

Query Match 1.5%; Score 10; DB 2; Length 777;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 OGGSTRTQOL 76
|||||
Db 230 OGGSTRTQOL 239

RESULT 32
D64138
penicillin-binding protein 1B homolog - *Haemophilus influenzae* (strain Rd KW20)
C:Species: *Haemophilus influenzae*
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 21-Jul-2000
C:Accession: D64138
R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Feldman
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
A:Reference number: A6400; MUID:95350630
A:Accession: D64138
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-781 <TIG>
A:Cross-references: GB:U52845; GB:I42023; NID:g3212236; PIDN:AAC23371.1; PID:g1574581
C:Superfamily: penicillin-binding protein 1B

Query Match 1.5%; Score 10; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 OGGSTRTQOL 76
|||||
Db 218 OGGSTRTQOL 227

RESULT 33
AC0412
penicillin-binding protein 1B (EC 2.4.2.-) [imported] - *Yersinia pestis* (strain CO92)
C:Species: *Yersinia pestis*
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C:Accession: AC0412
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
ll, M.; Rutherford, K.; Simmonds, T.; Skelton, J.; Stevens, K.; Whitehead, S.; Barril
Nature 413, 523-527, 2001
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0412
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-824 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC92623.1; PID:g15981319; GSPDB:GN00175
C:Genetics:
A:Gene: mrcB
C:Keywords: glycosyltransferase; pentosyltransferase

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:46:45 ; Search time 13.57 Seconds
(without alignments)
1900.311 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666

Sequence: 1 K1YDNKQIADLGSERVY.....TQSSNTTPDQONQNPQAP 666

Scoring table:

GAPOP 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 10

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : SwissProt.40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	719	1	PPPA_STRPN
2	53	8.0	637	1	PPPA_STROR
3	12	1.8	760	1	PPPB_BUCAI
4	10	1.5	577	1	MPPA_NEUCR
5	10	1.5	777	1	PPPB_VIBCH
6	10	1.5	781	1	PPPB_HAEIN

ALIGNMENTS

RESULT 1
ID PPPA_STRPN STANDARD; PRT; 719 AA.
AC Q04707;
DT 01-OCT-1994 (Rel. 30, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1A) (Exported protein 2).
GN PONA OR EXP2 OR SP0369.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain.";
RL EMBL J. 11:3831-3836(1992).

[2]
RN SEQUENCE FROM N.A.
RP STRAIN-R6;
RX MEDLINE=92325042; PubMed=1624444;
RA Martin C., Briese T., Hakenbeck R.;
RT "Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae and Streptococcus oralis with high homology to Escherichia coli penicillin-binding proteins 1a and 1b.";
RL J. Bacteriol. 174:4517-4523(1992).

[3]
RN SEQUENCE FROM N.A.

RP STRAIN-TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D., Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J., Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D., Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D., Holtzapple E., Khouli H., Wolf A.M., Uitterback T.R., Hansen C.L., McDonald L.A., Feldblyum T.V., Angiolini S., Dickinson T., Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.";
RL Science 293:498-506(2001).

[4]
RN SEQUENCE OF 293-369 FROM N.A.

RP STRAIN-R6X;
RX MEDLINE=95020625; PubMed=7934910;

RA Pearce B.J., Yin Y.B., Masure H.R.;

RT "Genetic identification of exported proteins in Streptococcus pneumoniae.";
RL Mol. Microbiol. 9:1037-1050(1993).

CC -I- FUNCTION: CELL WALL FORMATION.

CC -I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.

CC -I- SUBCELLULAR LOCATION: Secreted.

CC -I- MISCELLANEOUS: THESE STRAINS ARE PENICILLIN-SENSITIVE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: X67873; CAA48073.1; -

DR EMBL: X67872; CAA48072.1; -

DR EMBL: M90527; AAA26956.1; -

DR EMBL: AE007349; AAK74536.1; -

DR PIR: S28038; S28038.

DR TIGR: SP0369; -

DR InterPro: IPR001264; Transglycosyl.

DR InterPro: IPR001460; Transpeptidase.

DR Pfam: PF00912; Transglycosyl. 1.

DR Pfam: PF00905; Transpeptidase; 1.

DR Prodom: PD001895; Transglycosyl. 1.

CC Peptidoglycan synthesis; Antibiotic resistance; Cell wall;

CC Multifunctional enzyme; Complete proteome.

CC ACT_SITE 370 370 ACYLATED BY PENICILLIN (BY SIMILARITY).

FT DOMAIN 658 683 SER-RICH.

FT VARIANT 124 124 T -> A (IN STRAIN R6).

FT VARIANT 386 386 V -> I (IN STRAIN 63915).

FT VARIANT 388 388 D -> E (IN STRAIN R6).

FT VARIANT 397 397 E -> K (IN STRAIN 63915).

FT VARIANT 523 523 M -> I (IN STRAIN 63915).

FT VARIANT 533 533 D -> E (IN STRAIN 45607).

FT VARIANT 540 540 T -> S (IN STRAINS 63915 AND R6).

FT VARIANT 657 657 N -> S (IN STRAINS 45607 AND R6).

SO SEQUENCE 719 AA; 79758 MW; 5BD397E83B4B3AA6 CRC64;

Query Match 100.0%; Score 666; DB 1; Length 719;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADLGSRRVNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLRN 60
 DB 54 KIYDNKNOLIADLGSRRVNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLRN 113
 QY 61 LOSNSTLQGGSTLTQOLIKLTYFSTSDOTISKRAQDAMLAIOLEKATKQELITYINK 120
 DB 114 LOSNSTLQGGSTLTQOLIKLTYFSTSDOTISKRAQDAMLAIOLEKATKQELITYINK 173
 QY 121 VYMSNNGYNGQTAQNTYNGDLNNLSLPOLALLAGMPQAPNOVDYSHPAADRNLVYL 180
 DB 174 VYMSNNGYNGQTAQNTYNGDLNNLSLPOLALLAGMPQAPNOVDYSHPAADRNLVYL 233
 QY 181 SEKNNOGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINQVEETGYNLLT 240
 DB 234 SEKNNOGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLYKEVINQVEETGYNLLT 293
 QY 241 TGMDEVYTNDOEAKHLMIDYNTDEVYANPDELQVASTIVDYSNGKVIAQLGARHQQSN 300
 DB 294 TGMDEVYTNDOEAKHLMIDYNTDEVYANPDELQVASTIVDYSNGKVIAQLGARHQQSN 353
 QY 301 VSRGYNQAVETNRDWSGTMKPTIDVAPALEYGYDSTATIVHDEPNYNGENTPVYNNMDR 360
 DB 354 VSRGYNQAVETNRDWSGTMKPTIDVAPALEYGYDSTATIVHDEPNYNGENTPVYNNMDR 413
 QY 361 GFYGNITLQYALQOSRNVAVETLNKVGINRAKTFPLNGLIDYPSIHSNAISSNTTESD 420
 DB 414 GFYGNITLQYALQOSRNVAVETLNKVGINRAKTFPLNGLIDYPSIHSNAISSNTTESD 473
 QY 421 KKYGASSEKMAAAYAAPANGGYTKKPMYIHKVYFSDGSEKESNVGTRAKETTYAMTD 480
 DB 474 KKYGASSEKMAAAYAAPANGGYTKKPMYIHKVYFSDGSEKESNVGTRAKETTYAMTD 533
 QY 481 MKKTVLYTGGRAYLAMLPOAGKTGTSNTDEIENHITKTSOFVAPDELFAGYTRKYSM 540
 DB 534 MKKTVLYTGGRAYLAMLPOAGKTGTSNTDEIENHITKTSOFVAPDELFAGYTRKYSM 593
 QY 541 AWMTGYSNRLTPLVGNGLTVAAKVYRSMYTLSEGSNPEDMNIPBGLYRNGEYVFRNGAR 600
 DB 594 AWMTGYSNRLTPLVGNGLTVAAKVYRSMYTLSEGSNPEDMNIPBGLYRNGEYVFRNGAR 653
 QY 601 STWNSAPAPQPPSTESSSSSSSSSTQSSSTPSTNNSTTTNNNTQOOSNTTTPDOONON 660
 DB 654 STWNSAPAPQPPSTESSSSSSSSSTQSSSTPSTNNSTTTNNNTQOOSNTTTPDOONON 713
 QY 661 POPAP 666.
 DB 714 POPAP 719

RESULT 2
 PPBP_STROR STANDARD; PRT: 637 AA.

AC 000573;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Penicillin-binding protein 1A (PBP-1A) (Fragment).
 GN POMN.
 OS Streptococcus oralis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92325042; PubMed=1624444;
 RA Martin C., Briese T., Hakenbeck R.;
 RT "Nucleotide sequences of genes encoding penicillin-binding proteins
 from Streptococcus pneumoniae and Streptococcus oralis with high
 homology to Escherichia coli penicillin-binding proteins Ia and Ib.";
 RT J. Bacteriol. 174:4517-4523(1992).
 CC -I- FUNCTION: CELL WALL FORMATION.

CC -I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC
 DR EMBL: M90528; AAA26958.1; -
 DR PIR: B42893; B42893.
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001695; Transglycosyl; 1.
 KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
 KM Multifunctional enzyme.
 FT ACT_SITE 371 371
 FT NON_TER 637 637
 FT SEQUENCE 637 AA; 70891 MW; A6D198BCEA603A63 CRC64;
 ACYLATED BY PENICILLIN (BY SIMILARITY).

Query Match 8.0%; Score 53; DB 1; Length 637;
 Best Local Similarity 100.0%; Pred. No. 2e-45;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 GKVIAQLGARHQQSNVSGINQAVETNRDWSGTMKPTIDVAPALEYGYDSTA 338
 DB 340 GKVIAQLGARHQQSNVSGINQAVETNRDWSGTMKPTIDVAPALEYGYDSTA 392

RESULT 3
 PPBP_BUCAI STANDARD; PRT: 760 AA.

ID PPBP_BUCAI
 AC P57296;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Penicillin-binding protein 1B (PBP-1b) (PPBP) (Murein polymerase)
 DE [Includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-)
 DE (peptidoglycan Tgase); Penicillin-sensitive transpeptidase
 DE (EC 3.4.-.-) (DD-transpeptidase)].
 GN MRCP OR BU200.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OKYO 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RT Buchnera sp. APS.";
 RT Nature 407:81-86(2000).
 RL Nature 407:81-86(2000).
 CC -I- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
 CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
 CC TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
 CC SUBUNITS) (BY SIMILARITY).
 CC -I- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -I- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSPEPTIDASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: AP001118; BAB12917.1;
 DR InterPro: IPR001264; Transglycosyl.
 DR Pfam: PF00912; Transglycosyl. 1.
 DR ProDom: PD001895; Transglycosyl. 1.
 KW Peptidoglycan synthetase; Cell wall; Transferase; Glycosyltransferase;
 KW Hydrolyase; Multifunctional enzyme; Transmembrane; Signal-anchor;
 KW Antibiotic resistance; Complete proteome.
 FT DOMAIN 1
 FT TRANSMEM 9 29
 FT SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 136 760
 FT DOMAIN 392 684
 FT ACT_SITE 451 451
 FT ACYLATED BY PENICILLIN (BY SIMILARITY).
 SQ SEQUENCE 760 AA; 88002 MW; 00B2C5B51F6947D3 CRC64;

 Query Match 1.8%; Score 12; DB 1; Length 760;
 Best Local Similarity 100.0%; Pred. No. 0.00098;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 OY 67 OGSTLTQOLIK 78
 DB 204 OGSTLTQOLIK 215

 RESULT 4
 MPPE_NEUCR STANDARD; PRT: 577 AA.
 AC P23955;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitochondrial processing peptidase alpha subunit, mitochondrial
 DE precursor (EC 3.4.24.64) (Alpha-MPP).
 OS MPP.
 ON Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE OF N-TERMINUS.
 RX MEHLIN-90277682; PubMed=2141023;
 RA Schneider H., Arretz M., Wachter E., Neupert W.;
 RT "Matrix processing peptidase of mitochondria. Structure-function
 RT relationships";
 RL J. Biol. Chem. 265:9881-9887(1990).
 CC -1- FUNCTION: THE MITOCHONDRIAL PROCESSING PROTEASE CLEAVES
 CC PRESEQUENCES FROM MITOCHONDRIAL PROTEIN PRECURSORS.
 CC -1- CATALYTIC ACTIVITY: Release of N-terminal transit peptides from
 CC precursor proteins imported into the mitochondrion, typically with
 CC Arg in position P2.
 CC -1- SUBUNIT: HETERODIMER OF ALPHA AND BETA SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- DOMAIN: APPEARS TO CONTAIN TWO DOMAINS OF APPROXIMATELY EQUAL SIZE
 CC WHICH ARE SEPARATED BY A LOOP-LIKE SEQUENCE.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
 CC INSULINASE FAMILY. DOES NOT SEEM TO HAVE A PROTEASE ACTIVITY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: J05484; AAA33597.1; ALT_SEQ.

DR PIR: A36442; A36442.
 DR HSSP: P31800; 1BE3.
 DR InterPro: IPR001431; Peptidase_M16.
 DR Pfam: PF00675; Peptidase_M16; 1.
 DR PROSITE: PS00143; INSULINASE; 1.
 KW Hydrolyase; Metalloprotease; Mitochondrion; Transit peptide.
 FT TRANSMEM 1
 FT TRANSMEM 36 577
 FT CHAIN 1 35
 FT MITOCHONDRIAL
 FT ALPHA SUBUNIT.
 FT SER-RICH.
 SQ SEQUENCE 577 AA; 63027 MW; 44098F146BBDD1 CRC64;

 Query Match 1.5%; Score 10; DB 1; Length 577;
 Best Local Similarity 100.0%; Pred. No. 0.081;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 OY 615 ESSSSSSSDS 624
 DB 286 ESSSSSSSDS 295

 RESULT 5
 PBPB_VIBCH STANDARD; PRT: 777 AA.
 ID PBPB_VIBCH
 AC Q9KUC0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Penicillin-binding protein 1B (PBP-1b) (PBP1b) (Murein polymerase)
 DE (includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-))
 DE (Peptidoglycan Tgase); Penicillin-sensitive transpeptidase
 DE (EC 3.4.2.-) (DD-transpeptidase)).
 DE MRCB OR POND OR VC0602.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEHLIN-20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae";
 RL Nature 406:477-483(2000).
 CC -1- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
 CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
 CC TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
 CC SUBUNITS) (BY SIMILARITY).
 CC -1- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
 CC similarity).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSPEPTIDASE FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC

```

DR EMBL: AE004145; AAF93769.1; -
DR TIGR: VC0602; -
DR InterPro: IPR001264; Transglycosyl.
DR Pfam: PF00912; Transglycosyl; 1.
DR ProDom: PD001895; Transglycosyl; 1.
DR Peptidoglycan synthetase; Cell wall; Transferase; Glycosyltransferase;
KM Hydrolyase; Multifunctional enzyme; Transmembrane; Inner membrane;
KW Signal-anchor; Antibiotic resistance; Complete proteome.
FT DOMAIN 1 30 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 31 52 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT DOMAIN 1 (POTENTIAL).
FT DOMAIN 53 777 PERIPLASMIC (POTENTIAL).
FT DOMAIN 162 334 TRANSGLYCOSYLASE.
FT DOMAIN 415 709 TRANSPEPTIDASE.
FT ACT_SITE 476 476 ACYLATED BY PENICILLIN (BY SIMILARITY).
SQ SEQUENCE 777 AA; 87323 MW; BCF2676EA2AF1C CRC64;

Query Match 1.5%; Score 10; DB 1; Length 777;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 QGGSTLTQOL 76
Db 230 QGGSTLTQOL 239

RESULT 6
PBPB_HAEIN
ID PBPB_HAEIN STANDARD: PRT; 781 AA.
AC P45345;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1B (PBP-1b) (PBP1b) (Murein polymerase)
DE [includes: Penicillin-insensitive transglycosylase (EC 2.4.2.-)
DE (EC 3.4.-.-) (DD-transpeptidase)].
DE MRCH OR PONB OR H11725.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd. ";
RL Science 269:496-512(1995).
CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSPEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNITS) (BY SIMILARITY).
CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Inner membrane (By similarity).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSPEPTIDASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32845; AAC23371.1; -
DR TIGR: H11725; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR ProDom: PD001895; Transglycosyl; 1.
DR Peptidoglycan synthetase; Cell wall; Transferase; Glycosyltransferase;
KW Hydrolyase; Multifunctional enzyme; Inner membrane;
KM Antibiotic resistance; Complete proteome.
FT DOMAIN 151 322 TRANSGLYCOSYLASE.
FT DOMAIN 415 702 TRANSPEPTIDASE.
FT ACT_SITE 466 466 ACYLATED BY PENICILLIN (BY SIMILARITY).
SQ SEQUENCE 781 AA; 88256 MW; 769CD5DE9A6E488E CRC64;

Query Match 1.5%; Score 10; DB 1; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 QGGSTLTQOL 76
Db 218 QGGSTLTQOL 227

```

Search completed: June 13, 2002, 08:50:38
Job time: 233 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:46:25 ; Search time 34.32 Seconds

(Without alignments)
3357.069 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666
1 KIYDNKNQIADLGSERFVN.....TQSNSTTPQONQNPQAP 666Sequence: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 10

Total number of hits satisfying chosen parameters: 81

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP nhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP rivirus: *
16: SP bacteriaph: *
17: SP archaeap: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	666	100.0	719	2	Q9REU0 streptococc
2	479	71.9	719	2	Q9REU9 streptococc
3	351	52.7	608	2	Q549S1 streptococc
4	351	52.7	719	2	Q9REU3 streptococc
5	274	41.1	398	2	Q9RTM6 streptococc
6	262	39.3	719	2	Q57114 streptococc
7	223	33.5	310	2	Q52739 streptococc
8	223	33.5	310	2	Q52741 streptococc
9	223	33.5	310	2	Q52742 streptococc
10	223	33.5	310	2	Q52743 streptococc
11	207	29.6	718	2	Q9REU7 streptococc
12	197	28.8	398	2	Q87102 streptococc
13	192	27.2	719	2	Q9WU11 streptococc
14	181	27.2	719	2	Q54946 streptococc
15	181	27.2	719	2	Q9WUW0 streptococc
16	181	27.2	719	2	Q9WUW0 streptococc

17	181	27.2	719	2	Q9RET6 streptococc
18	181	27.2	719	2	Q9RET4 streptococc
19	181	27.2	719	2	Q54947 streptococc
20	181	27.2	719	2	Q54948 streptococc
21	181	27.2	719	2	Q54949 streptococc
22	163	24.5	163	2	Q9ETQ2 streptococc
23	142	20.3	719	2	Q9RET8 streptococc
24	139	20.9	163	2	Q9EM43 streptococc
25	134	20.1	310	2	Q52738 streptococc
26	120	18.0	719	2	Q9RET5 streptococc
27	105	15.8	398	2	Q87103 streptococc
28	105	15.8	398	2	Q9RTM7 streptococc
29	105	15.8	398	2	Q87105 streptococc
30	98	14.7	189	2	Q54918 streptococc
31	95	14.3	310	2	Q52737 streptococc
32	85	12.8	398	2	Q9RTM5 streptococc
33	80	12.0	398	2	Q87108 streptococc
34	79	11.9	719	2	Q54950 streptococc
35	76	11.4	398	2	Q87107 streptococc
36	60	9.0	163	2	Q9EM42 streptococc
37	60	9.0	398	2	Q9RTM8 streptococc
38	54	8.1	727	2	Q9ETC4 streptococc
39	53	8.0	398	2	Q87104 streptococc
40	50	7.5	163	2	Q9EM44 streptococc
41	41	6.2	163	2	Q9EM41 streptococc
42	41	6.2	310	2	Q93N72 streptococc
43	38	5.7	527	2	Q70037 streptococc
44	32	4.8	728	2	Q9P2G7 streptococc
45	30	4.5	163	2	Q9EM45 streptococc
46	27	4.1	400	2	Q87106 streptococc
47	23	3.5	721	16	Q99YL1 streptococc
48	22	3.3	163	2	Q9EM40 streptococc
49	20	3.0	139	2	Q54919 streptococc
50	20	3.0	310	2	Q52744 streptococc
51	18	2.7	793	2	Q9EXM9 streptococc
52	14	2.1	664	16	Q9C123 streptococc
53	14	2.1	778	14	Q9EXN1 streptococc
54	11	1.7	196	2	Q66087 streptococc
55	11	1.7	533	2	Q70040 streptococc
56	11	1.7	801	16	Q9C1H4 streptococc
57	11	1.7	821	2	Q70038 streptococc
58	11	1.7	821	16	Q97NE4 streptococc
59	11	1.7	966	16	Q9K7X8 streptococc
60	10	1.5	118	5	Q9U1K5 streptococc
61	10	1.5	119	5	Q9V401 streptococc
62	10	1.5	161	4	Q9NW20 streptococc
63	10	1.5	161	4	Q9BVL1 streptococc
64	10	1.5	161	4	Q9NY07 streptococc
65	10	1.5	259	4	Q9H1F6 streptococc
66	10	1.5	261	4	Q9NZ31 streptococc
67	10	1.5	294	5	Q01796 streptococc
68	10	1.5	504	5	Q02254 streptococc
69	10	1.5	516	2	Q9RQ72 streptococc
70	10	1.5	580	2	Q9K0K9 streptococc
71	10	1.5	584	2	Q9K1J2 streptococc
72	10	1.5	693	2	Q9KK21 streptococc
73	10	1.5	693	16	Q97N74 streptococc
74	10	1.5	693	2	Q9KK36 streptococc
75	10	1.5	696	2	Q9KK32 streptococc
76	10	1.5	714	10	Q9PJ70 streptococc
77	10	1.5	744	16	Q92R95 streptococc
78	10	1.5	755	16	Q98GP0 streptococc
79	10	1.5	767	16	Q92T08 streptococc
80	10	1.5	774	16	Q9X6W0 streptococc
81	10	1.5	775	16	Q9CNH5 streptococc

ALIGNMENTS

RESULT 1
Q9REU0
PRELIMINARY; PRT; 719 AA.

AC Q9REU0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBP1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP 1261;
 RX MEDLINE-21432820; PubMed=11549185;
 RA Ferroni A.; Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
 penicillin-resistant clinical isolates of Streptococcus pneumoniae
 serotype 23F from the nasopharyngeal flora of children.";
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL; AF210745; AAF17255.1; -
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79758 MW; 5BD397E83B4B3A6 CRC64;

Query Match 100.0%; Score 666; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLN 60
 DB 54 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLN 113
 QY 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAQEAMLAIQLEKATKQELITYYINK 120
 DB 114 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAQEAMLAIQLEKATKQELITYYINK 173
 QY 121 VYMSNGNYGMOTRAQNYVYKDLNNLSLPOLALLAGMPQAPNOVDPSHPEAADRRNLV 180
 DB 174 VYMSNGNYGMOTRAQNYVYKDLNNLSLPOLALLAGMPQAPNOVDPSHPEAADRRNLV 233
 QY 181 SEMKNGYISAEQYKAVMPTITDGLQSLKSASNPAYMDNTLKEYINQVEETGYNLLT 240
 DB 234 SEMKNGYISAEQYKAVMPTITDGLQSLKSASNPAYMDNTLKEYINQVEETGYNLLT 293
 QY 241 TGMADVTTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDSNGKVIAQLGARHOSSN 300
 DB 294 TGMADVTTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDSNGKVIAQLGARHOSSN 353
 QY 301 VSEFGINQAVETNRDMSGTKPITDVAAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 360
 DB 354 VSEFGINQAVETNRDMSGTKPITDVAAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 413
 QY 361 GFEGNTITLOALQOQRNVAVETLNKVGINRAKTFNLGIGIDVPSIHYSNALISNTTESD 420
 DB 414 GFEGNTITLOALQOQRNVAVETLNKVGINRAKTFNLGIGIDVPSIHYSNALISNTTESD 473
 QY 421 KRYGASSEKMAAAYAAAFANGGTYKPMYIHKVVSOGSEKESNVGTRAMKETTAAYMMT 480
 DB 474 KRYGASSEKMAAAYAAAFANGGTYKPMYIHKVVSOGSEKESNVGTRAMKETTAAYMMT 533
 QY 481 MKKTYLTYGTGRNAYLAAMPQAGKTGTSNTTDEEINNHKTSGFVAPDELFAGYTRYKYSN 540
 DB 534 MKKTYLTYGTGRNAYLAAMPQAGKTGTSNTTDEEINNHKTSGFVAPDELFAGYTRYKYSN 593
 QY 541 AYWGTGSNLTPLVNGNLVAAKVYSMMTYISEGSPEDDMNIPEDLYRNGEVEFVNGAR 600
 DB 594 AYWGTGSNLTPLVNGNLVAAKVYSMMTYISEGSPEDDMNIPEDLYRNGEVEFVNGAR 653
 QY 601 STWNSPAPQPPSTESSSSSDSSSTPSQSSSTPTSTNNSTTTNPNNNTTQOOSNTTPOOQN 660

DB 654 STWNSPAPQPPSTESSSSSDSSSTPSQSSSTPTSTNNSTTTNPNNNTTQOOSNTTPOOQN 713
 QY 661 POPAP 666
 DB 714 POPAP 719

RESULT 2
 ID Q9RET9 PRELIMINARY; PRT; 719 AA.
 AC Q9RET9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBP1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP 1513;
 RX MEDLINE-21432820; PubMed=11549185;
 RA Ferroni A.; Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
 penicillin-resistant clinical isolates of Streptococcus pneumoniae
 serotype 23F from the nasopharyngeal flora of children.";
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL; AF210746; AAF17256.1; -
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79830 MW; 5F6776B8DEAE3840 CRC64;

Query Match 71.9%; Score 479; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 479; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLN 60
 DB 54 KIYDNKNQIADLGSERRNAQANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLN 113
 QY 61 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAQEAMLAIQLEKATKQELITYYINK 120
 DB 114 LOSNSLGGSTLTQOLIKLTYFSTSDQTSRKAQEAMLAIQLEKATKQELITYYINK 173
 QY 121 VYMSNGNYGMOTRAQNYVYKDLNNLSLPOLALLAGMPQAPNOVDPSHPEAADRRNLV 180
 DB 174 VYMSNGNYGMOTRAQNYVYKDLNNLSLPOLALLAGMPQAPNOVDPSHPEAADRRNLV 233
 QY 181 SEMKNGYISAEQYKAVMPTITDGLQSLKSASNPAYMDNTLKEYINQVEETGYNLLT 240
 DB 234 SEMKNGYISAEQYKAVMPTITDGLQSLKSASNPAYMDNTLKEYINQVEETGYNLLT 293
 QY 241 TGMADVTTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDSNGKVIAQLGARHOSSN 300
 DB 294 TGMADVTTNVDOEAQKHLMDIYNTDEVYAYPDELOVASTIVDSNGKVIAQLGARHOSSN 353
 QY 301 VSEFGINQAVETNRDMSGTKPITDVAAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 360
 DB 354 VSEFGINQAVETNRDMSGTKPITDVAAPALEYGYDSTATIVHDEPNYNGTNPVYNMNR 413
 QY 361 GFEGNTITLOALQOQRNVAVETLNKVGINRAKTFNLGIGIDVPSIHYSNALISNTTESD 420
 DB 414 GFEGNTITLOALQOQRNVAVETLNKVGINRAKTFNLGIGIDVPSIHYSNALISNTTESD 473
 QY 421 KRYGASSEKMAAAYAAAFANGGTYKPMYIHKVVSOGSEKESNVGTRAMKETTAAYMMT 479

Db 474 KKGASSEKMAAAYAFANGGTYYKPMYHKVPSDGESEFNVGTRAMETAYMMT 532

RESULT 3
ID 054951 PRELIMINARY; PRT: 608 AA.
AC 054951;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN Streptococcus pneumoniae.
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8250;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-resistant Streptococcus pneumoniae isolated in
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain.";
RL EMBO J. 11:3831-3836(1992).
DR EMBL; X67871; CAA48071.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam; PF00912; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR Prodom; PD001895; Transglycosyl; 1.
FT NON_TER 608
SQ SEQUENCE 608 AA; 67918 MW; 68F5C09E2AF9114 CRC64;

Query Match 52.7%; Score 351; DB 2; Length 608;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSERRVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFRLN 60
DB 54 KIYDNKNLIDLGSERRVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFRLN 113
QY 61 LQSNLSOGSSTLTQOLIKLTYFSTSDQTSRKAQEMALAIQLEOKATKQELITYYINK 120
DB 114 LQSNLSOGSSTLTQOLIKLTYFSTSDQTSRKAQEMALAIQLEOKATKQELITYYINK 173
QY 121 YVMSNGNTGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPPEAAQDRRLVL 180
DB 174 YVMSNGNTGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPPEAAQDRRLVL 233
QY 181 SEMKNQGYISAQYKAVNPTDGLQSLKSASNPAYMDNVLKVINOVEERTGYNLLT 240
DB 234 SEMKNQGYISAQYKAVNPTDGLQSLKSASNPAYMDNVLKVINOVEERTGYNLLT 293
QY 241 TGMVYTVNDQEAQKHLMDIYNTDEVYVAPDDELQVASTIVDVNSGKYTAQLGARHSSN 300
DB 294 TGMVYTVNDQEAQKHLMDIYNTDEVYVAPDDELQVASTIVDVNSGKYTAQLGARHSSN 353
QY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGT 351
DB 354 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGT 404

RESULT 4
ID 09RET3 PRELIMINARY; PRT: 719 AA.
AC 09RET3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A.
GN BPB1A.

OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM 4200;
RX MEDLINE=21432820; PubMed=11549185;
RA Ferriol A., Berche P.;
RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
RT penicillin-resistant clinical isolates of Streptococcus pneumoniae
RT serotype 23F from the nasopharyngeal flora of children.";
RL J. Med. Microbiol. 50:828-832(2001).
DR EMBL; AF210752; AAF17262.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam; PF00912; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR Prodom; PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79781 MW; DB5993f6BD47F72D CRC64;

Query Match 52.7%; Score 351; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNLIDLGSERRVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFRLN 60
DB 54 KIYDNKNLIDLGSERRVNAQANDIPDYLKAIYSIEDHREFDRGIDTIRILGAFRLN 113
QY 61 LQSNLSOGSSTLTQOLIKLTYFSTSDQTSRKAQEMALAIQLEOKATKQELITYYINK 120
DB 114 LQSNLSOGSSTLTQOLIKLTYFSTSDQTSRKAQEMALAIQLEOKATKQELITYYINK 173
QY 121 YVMSNGNTGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPPEAAQDRRLVL 180
DB 174 YVMSNGNTGMOTAAQNYGKDLNLSLPOLALLAGMPQAPNOYDPSHPPEAAQDRRLVL 233
QY 181 SEMKNQGYISAQYKAVNPTDGLQSLKSASNPAYMDNVLKVINOVEERTGYNLLT 240
DB 234 SEMKNQGYISAQYKAVNPTDGLQSLKSASNPAYMDNVLKVINOVEERTGYNLLT 293
QY 241 TGMVYTVNDQEAQKHLMDIYNTDEVYVAPDDELQVASTIVDVNSGKYTAQLGARHSSN 300
DB 294 TGMVYTVNDQEAQKHLMDIYNTDEVYVAPDDELQVASTIVDVNSGKYTAQLGARHSSN 353
QY 301 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGT 351
DB 354 VSFGINQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIVHDEPYNPGT 404

RESULT 5
ID 09R7M6 PRELIMINARY; PRT: 398 AA.
AC 09R7M6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN BPB1A.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=#10/219;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
RT of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).

DR EMBL: AB006874; BAA32070.1; -
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 398
 SQ SEQUENCE 398 AA; 43613 MW; D2C0E784E49E3451 CRC64;

Query Match 41.1%; Score 274; DB 2; Length 398;
 Best Local Similarity 99.7%; Pred. No. 3.9e-270;
 Matches 374; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 267 VAYPDELQVASTIVDSNGKVTALQAGRHSSNVSGFNGINQAVETNRDMSGTMRKPTTDYA 326
 DB 1 VAYPDELQVASTIVDSNGKVTALQAGRHSSNVSGFNGINQAVETNRDMSGTMRKPTTDYA 60
 QY 327 PALEGVYSTATIVHDEPNYNGTPTVYNNDRGFGFNGITLQALQOQRNPAVETLNK 386
 DB 61 PALEGVYSTATIVHDEPNYNGTPTVYNNDRGFGFNGITLQALQOQRNPAVETLNK 120
 QY 387 VGLNRAKTEFLNGIDYPSIHYSNALSSNTTESDKKYGASSEKMAAAYAAFAFGCTYKRP 446
 DB 121 VGLNRAKTEFLNGIDYPSIHYSNALSSNTTESDKKYGASSEKMAAAYAAFAFGCTYKRP 180
 QY 447 MTIHKVYSDGSEKFSNVGTGRAMKETAYMMTDMKTVLTGTGRNAVILAMPQAGKTG 506
 DB 181 MTIHKVYSDGSEKFSNVGTGRAMKETAYMMTDMKTVLTGTGRNAVILAMPQAGKTG 240
 QY 507 TSNYTDDELENIHKTQFOFAPDELFGYTRKYSMAVWTGYSNLTFLVNGSLTVAKAYR 566
 DB 241 TSNYTDDELENIHKTQFOFAPDELFGYTRKYSMAVWTGYSNLTFLVNGSLTVAKAYR 300
 QY 567 SMWYTLSEGSNPEDMNIPEGLYRNGEFVKNGARSTWNSPAPQPPSTSSSSSDSSTS 626
 DB 301 SMWYTLSEGSNPEDMNIPEGLYRNGEFVKNGARSTWNSPAPQPPSTSSSSSDSSTS 360
 QY 627 QSSSTPTSTNNSTTT 641
 DB 361 QSSSTPTSTNNSTTT 375

RESULT 6
 Q57114 PRELIMINARY; PRT; 719 AA.
 ID 057114
 AC 057114;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RX MEDLINE=96012191; PubMed=7574521;
 RA Coffey T.J., Daniels M., McDougal L.K., Dowson C.G., Tenover F.C., Spratt B.G.;
 RA "Genetic analysis of clinical isolates of Streptococcus pneumoniae with high-level resistance to expanded-spectrum cephalosporins";
 RL Antimicrob. Agents Chemother. 39:1306-1313(1995).
 DR EMBL: 249095; CAA88918.1; -
 DR EMBL: 249094; CAA88917.1; -
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR Prodom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79712 MW; 5396607C0E67D06E CRC64;

Query Match 39.3%; Score 262; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 1.1e-257;
 Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKNQIADIGSERRNAQANDIPTDVKAVIASEIHRFPDHDGIDTIRILGAFNR 60
 DB 54 KIYDNKNQIADIGSERRNAQANDIPTDVKAVIASEIHRFPDHDGIDTIRILGAFNR 113
 QY 61 LOSNSLOGSTTLQOLIKLTYFSTSDQTSKRAQEWALAIQEQKATKQELITYINK 120
 DB 114 LOSNSLOGSTTLQOLIKLTYFSTSDQTSKRAQEWALAIQEQKATKQELITYINK 173
 QY 121 VYMSNGYGMQTAQNYGKDLNNLSLPLALLAGMPQAPNOYDPYSHPEAADRRLVYL 180
 DB 174 VYMSNGYGMQTAQNYGKDLNNLSLPLALLAGMPQAPNOYDPYSHPEAADRRLVYL 233
 QY 181 SEKNNGYISAEQYERAVNPTIDGLOSLSASNTAYMDNYLKEYINOVEETGYNLLT 240
 DB 234 SEKNNGYISAEQYERAVNPTIDGLOSLSASNTAYMDNYLKEYINOVEETGYNLLT 293
 QY 241 TGMDEVYTNVDOEAQKHLMDIYN 262
 DB 294 TGMDEVYTNVDOEAQKHLMDIYN 315

RESULT 7
 O52739 PRELIMINARY; PRT; 310 AA.
 ID 052739
 AC 052739;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 GN PONA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RX MEDLINE=96287565; PubMed=9624469;
 RA Smith A.M., Klugman K.P.;
 RA "Alterations in PBPIA essential for high-level penicillin resistance in Streptococcus pneumoniae";
 RL Antimicrob. Agents Chemother. 42:1329-1333(1998).
 DR EMBL: AF046233; AAC24698.1; -
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 310
 SQ SEQUENCE 310 AA; 34445 MW; 31C692D48789576D CRC64;

Query Match 33.5%; Score 223; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 2.7e-218;
 Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LMDIYNTDEVAYPDELQVASTIVDSNGKVTALQAGRHSSNVSGFNGINQAVETNRDMSG 316
 DB 1 LMDIYNTDEVAYPDELQVASTIVDSNGKVTALQAGRHSSNVSGFNGINQAVETNRDMSG 60
 QY 317 STMKPPTDPALEGVYSTATIVHDEPNYNGTPTVYNNDRGFGFNGITLQALQOQR 376
 DB 61 STMKPPTDPALEGVYSTATIVHDEPNYNGTPTVYNNDRGFGFNGITLQALQOQR 120
 QY 377 NYPAVETLNKVGINRAKTEFLNGIDYPSIHYSNALSSNTTESDKKYGASSEKMAAAYAA 436
 DB 121 NYPAVETLNKVGINRAKTEFLNGIDYPSIHYSNALSSNTTESDKKYGASSEKMAAAYAA 180
 QY 437 FANGGTYYKPMYIHKVYFSDGSEKFSNVGTGRAMKETAYMMT 479
 DB 181 FANGGTYYKPMYIHKVYFSDGSEKFSNVGTGRAMKETAYMMT 223

RESULT	8			
052741				
ID	052741	PRELIMINARY;	PRT;	310 AA.
AC	052741;			
DT	01-JUN-1998	(TREMBLrel. 06, Created)		
DT	01-JUN-1998	(TREMBLrel. 06, last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).			
CN	PONA.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1313;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP				
RC	STRAIN=7851;			
RC	MEDLINE=98287565; PubMed=9624469;			
RA	Smith A.M., Klugman K.P.;			
RT	"Alterations in PBP 1A essential for high-level penicillin resistance			
RT	in Streptococcus pneumoniae.";			
RL	Antimicrob. Agents Chemother. 42:1329-1333(1998).			
DR	EMBL; AF046235; AAC24700.1; "			
DR	InterPro; IPR001460; Transpeptidase.			
DR	Pfam; PF00905; Transpeptidase; 1.			
FT	NON_TER	1		
FT	NON_TER	1		
FT	NON_TER	310		
FT	NON_TER	310		
SEQUENCE	310 AA;	34433 MW;	6E2EDFA50AB28CF3	CRC64;

Query Match	33.5%	Score 223	DB 2	Length 310
Best Local Similarity	100.0%	Pred. NO.	2.7e-218	
Matches 223	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	257	LMIDYINDEYAYADDELOVASTIVDVSNKGVINQOLCAROSSNVSGINQOAVETNDDWG	31.6
Db	1	LMIDYINDEYAYADDELOVASTIVDVSNKGVINQOLCAROSSNVSGINQOAVETNDDWG	60
Qy	317	STMKPTDYAPALEYGYVDSTATIVHDEPYNYPCTNTPVYNMDDRGYGNITTLQYALQOSR	376
Db	61	STMKPTDYAPALEYGYVDSTATIVHDEPYNYPCTNTPVYNMDDRGYGNITTLQYALQOSR	120
Qy	377	NVPAYETLNKYGILNRAKTFELNGCIDYPSLIHYSNAISSNTTESDKKYGAASSEKMAAAYAA	436
Db	121	NVPAYETLNKYGILNRAKTFELNGCIDYPSLIHYSNAISSNTTESDKKYGAASSEKMAAAYAA	180
Qy	437	FANGCTYYKPMYIHKYVESDSEKEEFSNVGIRAKKETTAIYMT	479
Db	181	FANGCTYYKPMYIHKYVESDSEKEEFSNVGIRAKKETTAIYMT	223

RESULT	9			
052742				
ID	052742	PRELIMINARY;	PRT;	310 AA.
AC	052742;			
DT	01-JUN-1998	(TREMBLrel. 06, Created)		
DT	01-JUN-1998	(TREMBLrel. 06, last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).			
GN	PONA.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus.			
OX	NCHI_TaxID=1313;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NG4;			
RX	MEDLINE=98287565; Pubmed=9624469;			
RA	Smith A.M., Klugman K.P.:			
RT	"Alterations in PBP 1A essential for high-level penicillin resistance			
RL	in Streptococcus pneumoniae.";			
RT	Antimicrob. Agents Chemother. 42:1329-1333(1998).			

DR EMBL: AF046236; AAC24701.1; -.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00905; Transpeptidase; 1.
FT NON_TER 1 1
FT NON_TER 310 310
SQ SEQUENCE 310 AA; 34480 MW; 39649X7387C5A56D CRC64;

Query Match	33.5%	Score 223	DB 2	length 310
Best Local Similarity	100.0%	Pred. No.	2.7e-218	
Matches 223	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	317	STMKPIIDYAPALEYGYVDSYDSTATIVHDEPYNYPGTNTPYVNMWDRGFGNITLQYALQOQR	376
Db	61	STMKPIIDYAPALEYGYVDSYDSTATIVHDEPYNYPGTNTPYVNMWDRGFGNITLQYALQOQR	120
QY	257	LMIDIYNDEYAYAPDDELOQYASTIVDYSNGVIAIQLGARQSSNVSGIQAQVETNRDMG	316
Db	1	LMIDIYNDEYAYAPDDELOQYASTIVDYSNGVIAIQLGARQSSNVSGIQAQVETNRDMG	60
QY	377	NVPAYEVLINKVGLNRAKTFLLGCLIDYPSIHSNAISSNTTESDKKYGASSEKMAAAYAA	436
Db	121	NVPAYEVLINKVGLNRAKTFLLGCLIDYPSIHSNAISSNTTESDKKYGASSEKMAAAYAA	180
QY	437	FANGGYKKPMYIHKVYESOSSEKEEFSNVCGRRAKETTAAVMT	479
Db	181	FANGGYKKPMYIHKVYESOSSEKEEFSNVCGRRAKETTAAVMT	223

RESULT 10

ID	053743	PRELIMINARY;	PRT;	310 AA.
DT	01-JUN-1998	(TREMBLrel_06, Created)		
DT	01-JUN-1998	(TREMBLrel_16, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel_09, Last annotation update)		
DE	PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).			
GN	PONA.			
OS	Streptococcus pneumoniae.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_Taxid=1313;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=I7619;			
RX	MEDLINE=98287565; PubMed=9624469;			
RA	Smith A.M., Klueman K.P.;			
RT	"Alterations in PBP 1A essential for high-level penicillin resistance			
RT	in Streptococcus pneumoniae."			
RL	Animicrob. Agents Chemother. 42:1329-1333(1998).			
DR	EMBL; AF0406237; AAC24702.1; -.			
DR	InterPro; IPR001460; Transpeptidase.			
DR	Pfam; PF00905; Transpeptidase; 1.			
FT	NON_TER	1		
FT		1		
FT		310		
SQ	SEQUENCE	310 AA; 34461 MW; 6E280216D7528CF3 CRC64;		

Query Match	33.5%	Score 223	DB 2	Length 310
Best Local Similarly	100.0%	Pred. No.	2.7e-218	
Matches 223	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	257	LMYINNDDEVAAYAPDDELOVAASTVIVDSNCKVIAOLGARHOSSNSFGINQAVETNRDM	316
Db	1	LMQIDYNDEVAAYAPDDELOVAASTVIVDSNCKVIAOLGARHOSSNSFGINQAVETNRDM	60
Qy	317	STMKPITDIYAPALEYGYIDSTATVHDEPYNINYGCTNTPVYNMDRGYFGNTLLOYALQOOR	376
Db	61	STMKPITDIYAPALEYGYIDSTATVHDEPYNINYGCTNTPVYNMDRGYFGNTLLOYALQOOR	120
Qy	377	NVPAPVETLNNKGLNRRATFLNGLQIDVPSIHYSNALSSNTTESDKRYGASSEKMAAAYAA	436
Db	121	NVPAPVETLNNKGLNRRATFLNGLQIDVPSIHYSNALSSNTTESDKRYGASSEKMAAAYAA	180

QY 437 FANGGTYRKPMYIHKVFSDSSEKFSNVGTRAMKETAYMMT 479
 DB 181 FANGGTYRKPMYIHKVFSDSSEKFSNVGTRAMKETAYMMT 223

RESULT 11

052740 PRELIMINARY: PRT: 310 AA.

ID 052740;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

RT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).

GN PONA.

OS Streptococcus pneumoniae.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1313;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=56739;

RX MEDLINE=98287565; PubMed=9624469;

RA Smith A.M., Klugman K.P.; "Alterations in PBP 1A essential for high-level penicillin resistance

RT "Alterations in PBP 1A essential for high-level penicillin resistance

RT Antimicrob. Agents Chemother. 42:1329-1333(1998).

DR EMBL; AF046234; AAC24699.1; -

DR InterPro: IPR001460; Transpeptidase.

DR Pfam: PF00905; Transpeptidase; 1.

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

RA Ferroni A., Berche P.; "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst

RT penicillin-resistant clinical isolates of Streptococcus pneumoniae

RT serotype 23F from the nasopharyngeal flora of children.";

RL J. Med. Microbiol. 50:828-832(2001).

DR EMBL; AF210748; AAF17258.1; -

DR InterPro: IPR001264; Transglycosyl.

DR InterPro: IPR001460; Transpeptidase.

DR Pfam: PF00912; Transglycosyl; 1.

DR Pfam: PF00905; Transpeptidase; 1.

DR Prodom: PD001895; Transglycosyl; 1.

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

FT NON_TER

Query Match 28.8%; Score 192; DB 2; Length 398;
 Best Local Similarity 100.0%; Pred. No. 1.3e-186;
 Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 VAYPDELQVASTIVDSNGKYIAQLGARHSSNVSEFGINQAVETNRDMSGTMRPTDVA 326
 DB 1 VAYPDELQVASTIVDSNGKYIAQLGARHSSNVSEFGINQAVETNRDMSGTMRPTDVA 60

```

QY 327 PALEYGYDSTATIVHDEPNYPTGNTPEYNNMDRGYFNGITLOALQOSRNPVAVETLTK 386
Db 61 PALEYGYDSTATIVHDEPNYPTGNTPEYNNMDRGYFNGITLOALQOSRNPVAVETLTK 120
QY 387 VGLNRAKTFNLGLGIDYPSIHYNSNAISSNTTESDSKKYASSEKKAAYAAFAANGCTYYKP 446
Db 121 VGLNRAKTFNLGLGIDYPSIHYNSNAISSNTTESDSKKYASSEKKAAYAAFAANGCTYYKP 180
QY 447 MYHKRVFSDGS 458
Db 181 MYHKRVFSDGS 192

RESULT 14
O9WM11 PRELIMINARY: PRT: 719 AA.
AC O9WM11;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PERICILLIN-BINDING PROTEIN 1A.
GN BPPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=URU-E159, M134, URU-E135, AND URU-E157;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 0:0-0(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=URU-E206;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 145:0-0(1999).
DR EMBL: AF139890; AAD43073.1; -
DR EMBL: AF139887; AAD43070.1; -
DR EMBL: AF139888; AAD43071.1; -
DR EMBL: AF139889; AAD43072.1; -
DR EMBL: AF159448; AAD48430.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
SQ SEQUENCE 719 AA; 79701 MW; 3EDC8A21C94DAC97 CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 IRIIGAFLRNLSNSLOGSSTLTQOLIKLTFYSTSTSDQITSRKAQEAFLAIOLEQKATK 110
Db 104 IRIIGAFLRNLSNSLOGSSTLTQOLIKLTFYSTSTSDQITSRKAQEAFLAIOLEQKATK 163
QY 111 QELITYINKYMSNGNGMGTAAQNTYKGLNLSLPQALLAGMPQAPQYDPYSHE 170
Db 164 QELITYINKYMSNGNGMGTAAQNTYKGLNLSLPQALLAGMPQAPQYDPYSHE 223
QY 171 AAQRRNLVISEMKNQGTISEQYEKAVNPITDGLQSLKASNPAYMDNYLKEVINOV 230
Db 224 AAQRRNLVISEMKNQGTISEQYEKAVNPITDGLQSLKASNPAYMDNYLKEVINOV 283
QY 231 E 231

```

```

Db 284 E 284

RESULT 15
O54946 PRELIMINARY: PRT: 719 AA.
AC O54946;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE PERICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA OR BPPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=456;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain.";
RL EMBO J. 11:3831-3836(1992).
RN [2]
RP SEQUENCE OF 310-619 FROM N.A.
RC STRAIN=8303, AND 35193;
RA Smith A.M., Klugman K.P.;
RL submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 320-717 FROM N.A.
RC STRAIN=#17/246, #20/B98, AND #27/SHA3;
RX MEDLINE=96409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a thr-371 substitution in a conserved amino acid motif
RT of penicillin-binding protein 1A with penicillin resistance of
RT streptococcus pneumoniae";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-665;
RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V
RT clone of Streptococcus pneumoniae arose by large recombinational
RT replacements of the cpsA-pbpA region.";
RL Microbiology 0:0-0(1999).
DR EMBL: X67868; CAA48068.1; -
DR EMBL: AF046230; AAC24695.1; -
DR EMBL: AB006876; BAA32072.1; -
DR EMBL: AF139883; AAD43066.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl. 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA; 79736 MW; F090F91522C834DC CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 IRIIGAFLRNLSNSLOGSSTLTQOLIKLTFYSTSTSDQITSRKAQEAFLAIOLEQKATK 110
Db 104 IRIIGAFLRNLSNSLOGSSTLTQOLIKLTFYSTSTSDQITSRKAQEAFLAIOLEQKATK 163
QY 111 QELITYINKYMSNGNGMGTAAQNTYKGLNLSLPQALLAGMPQAPQYDPYSHE 170
Db 164 QELITYINKYMSNGNGMGTAAQNTYKGLNLSLPQALLAGMPQAPQYDPYSHE 223
QY 231 E 231

```

OY 171 AADRRNLVLESEKNGYISAQYERKAVNPTITDGLQSLKSASNYPAYMDNLYKEVINOV 230
 DB 224 AADRRNLVLESEKNGYISAQYERKAVNPTITDGLQSLKSASNYPAYMDNLYKEVINOV 283
 OY 231 E 231
 DB 284 E 284

RESULT 16

O9WVW0 PRELIMINARY: PRT: 719 AA.
 AC O9WVW0;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PO-342, PO-273, AND PO-341;
 RA Coffey T.J., Daniels M., Enright M.C., Spratt B.G.;
 RT "Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of Streptococcus pneumoniae arose by large recombinational replacements of the cpsA-pbpA region."
 RL Microbiology 0:0-0(1999).
 DR EMBL; AF139886; AADA3069.1; -;
 DR EMBL; AF139884; AADA3067.1; -;
 DR EMBL; AF139885; AADA3068.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79622 MW; 7FBA6A75EA8FF8B3 CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;

Best Local Similarity 100.0%; Pred. No. 3.6e-175;

Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 IRLGAFLENLQSNLSQSGSTLTQOLIKLTFSTSDQTSRKAQEWLAIOLEQKATK 110
 DB 104 IRLGAFLENLQSNLSQSGSTLTQOLIKLTFSTSDQTSRKAQEWLAIOLEQKATK 163
 OY 111 QELITYYINKVYMSNGNYGMQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHP 170
 DB 164 QELITYYINKVYMSNGNYGMQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHP 223
 OY 171 AADRRNLVLESEKNGYISAQYERKAVNPTITDGLQSLKSASNYPAYMDNLYKEVINOV 230
 DB 224 AADRRNLVLESEKNGYISAQYERKAVNPTITDGLQSLKSASNYPAYMDNLYKEVINOV 283
 OY 231 E 231
 DB 284 E 284

RESULT 17

O9RET6 PRELIMINARY: PRT: 719 AA.
 AC O9RET6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP 1053; PubMed-11549185;
 RX MEDLINE-21432820; PubMed-11549185;
 RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children."
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL; AF210749; AAF17259.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79638 MW; 1DAE3C5937048E16 CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;

Best Local Similarity 100.0%; Pred. No. 3.6e-175;

Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 IRLGAFLENLQSNLSQSGSTLTQOLIKLTFSTSDQTSRKAQEWLAIOLEQKATK 110
 DB 104 IRLGAFLENLQSNLSQSGSTLTQOLIKLTFSTSDQTSRKAQEWLAIOLEQKATK 163
 OY 111 QELITYYINKVYMSNGNYGMQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHP 170
 DB 164 QELITYYINKVYMSNGNYGMQTAQNYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHP 223
 OY 171 AADRRNLVLESEKNGYISAQYERKAVNPTITDGLQSLKSASNYPAYMDNLYKEVINOV 230
 DB 224 AADRRNLVLESEKNGYISAQYERKAVNPTITDGLQSLKSASNYPAYMDNLYKEVINOV 283
 OY 231 E 231
 DB 284 E 284

RESULT 18

O9RET4 PRELIMINARY: PRT: 719 AA.
 AC O9RET4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP 22861;
 RX MEDLINE-21432820; PubMed-11549185;
 RA Ferroni A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst penicillin-resistant clinical isolates of Streptococcus pneumoniae serotype 23F from the nasopharyngeal flora of children."
 RL J. Med. Microbiol. 50:828-832(2001).
 DR EMBL; AF210751; AAF17261.1; -;
 DR InterPro: IPR001264; Transglycosyl.
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00912; Transglycosyl; 1.
 DR Pfam: PF00905; Transpeptidase; 1.
 DR ProDom: PD001895; Transglycosyl; 1.
 SQ SEQUENCE 719 AA; 79567 MW; 7B6EDCCBB8BCF286 CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;

Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 51 IRIILAFILNLSNLSGGSTLTQOLIKITFYSTSDOTISRKAQEAWLAIQLEOKATK 110
DB 104 IRIILAFILNLSNLSGGSTLTQOLIKITFYSTSDOTISRKAQEAWLAIQLEOKATK 163
OY 111 OEILFYIINKVYMSNGNYGMQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 170
DB 164 OEILFYIINKVYMSNGNYGMQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 223
OY 171 AAODRRNLVLSEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQV 230
DB 224 AAODRRNLVLSEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQV 283
OY 231 E 231
DB 284 E 284

RESULT 19
O54947 PRELIMINARY: PRT: 719 AA.
ID 054947
AC 054947:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DF 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=681;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain.";
RL EMBL J. 11:3831-3836 (1992).
DR EMBL; X67867; CAA48067.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719
SQ SEQUENCE 719 AA; 79710 MW; D90EBA2D5380D9E1 CRC64;
```

Query Match 27.2%; Score 181; DB 2; Length 719;

Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 51 IRIILAFILNLSNLSGGSTLTQOLIKITFYSTSDOTISRKAQEAWLAIQLEOKATK 110
DB 104 IRIILAFILNLSNLSGGSTLTQOLIKITFYSTSDOTISRKAQEAWLAIQLEOKATK 163
OY 111 OEILFYIINKVYMSNGNYGMQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 170
DB 164 OEILFYIINKVYMSNGNYGMQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 223
OY 171 AAODRRNLVLSEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQV 230
DB 224 AAODRRNLVLSEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQV 283
OY 231 E 231
DB 284 E 284
```

RESULT 20

```

O54948 PRELIMINARY: PRT: 719 AA.
ID 054948
AC 054948:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DF 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=670;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
RT South Africa and Spain.";
RL EMBL J. 11:3831-3836 (1992).
DR EMBL; X67867; CAA48067.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 719
SQ SEQUENCE 719 AA; 79611 MW; 1DBFDDA93704930B CRC64;
```

Query Match 27.2%; Score 181; DB 2; Length 719;

Best Local Similarity 100.0%; Pred. No. 3.6e-175;
Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 51 IRIILAFILNLSNLSGGSTLTQOLIKITFYSTSDOTISRKAQEAWLAIQLEOKATK 110
DB 104 IRIILAFILNLSNLSGGSTLTQOLIKITFYSTSDOTISRKAQEAWLAIQLEOKATK 163
OY 111 OEILFYIINKVYMSNGNYGMQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 170
DB 164 OEILFYIINKVYMSNGNYGMQTAQNYGKDLNLSLPOLALLAGMPQAPNOYDPYSHE 223
OY 171 AAODRRNLVLSEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQV 230
DB 224 AAODRRNLVLSEMKNGYISAEQYKAVNPTITDGLQSLKSASNPAYMDNLTKEVINQV 283
OY 231 E 231
DB 284 E 284
```

RESULT 21

```

O54949 PRELIMINARY: PRT: 719 AA.
ID 054949
AC 054949:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DF 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56742;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
RT clones of penicillin-resistant Streptococcus pneumoniae isolated in
```

RT South Africa and Spain.";
 RL EMBL J. 11:3831-3836(1992).
 DR EMBL; X67869; CAA48069.1; -;
 DR InterPro; IPR001264; Transglycosyl.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00905; Transglycosyl; 1.
 DR Pfam; PF00905; Transpeptidase; 1.
 DR ProDom; PD001895; Transglycosyl; 1.
 FT NON_TER 719 719
 SO SEQUENCE 719 AA; 79750 MW; 4B9B6CE0CB861CD CRC64;

Query Match 27.2%; Score 181; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 3.6e-175;
 Matches 181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 51 IRLGAFRLNLSLOGGTLTQOLIKLTFSTSDQTSRKAQEWLAIOLEQATK 110
 DB 104 IRLGAFRLNLSLOGGTLTQOLIKLTFSTSDQTSRKAQEWLAIOLEQATK 163
 OY 111 QEILFYINKVYNSNGNYGQTAQNYGKDLNLSLPOLALAGMPQAPNOYDPSHPE 170
 DB 164 QEILFYINKVYNSNGNYGQTAQNYGKDLNLSLPOLALAGMPQAPNOYDPSHPE 223
 OY 171 AADRRNLVSEKNGYISAEYKAVNPTIDGLOSLKASANNYPAYMDNTEKEVINOV 230
 DB 224 AADRRNLVSEKNGYISAEYKAVNPTIDGLOSLKASANNYPAYMDNTEKEVINOV 283
 OY 231 E 231
 DB 284 E 284

RESULT 22
 O9E02 PRELIMINARY; PRT; 163 AA.
 AC O9E02:
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-950423; AND 960036;
 RA Overweg K., Bogaert D., Sluifster M., de Groot R., Hermans P.W.M.;
 RT "Molecular characterization of Streptococcus pneumoniae penicillin-
 resistance in the Netherlands.";
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ403978; CAC20960.1; -;
 DR EMBL; AJ403975; CAC20957.1; -;
 FT NON_TER 163 163
 FT NON_TER 163 163
 SO SEQUENCE 163 AA; 18477 MW; 99C44911EB96E5D3 CRC64;

Query Match 24.5%; Score 163; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 2e-157;
 Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 213 SNYPAYMDNTEKEVINOVSEETGYNLLTGMVYTNVDOEAKHMDIYNTDEYVYPPD 272
 DB 1 SNYPAYMDNTEKEVINOVSEETGYNLLTGMVYTNVDOEAKHMDIYNTDEYVYPPD 60
 OY 273 ELQVASTIVDSNGKVIAGLARGHSSNVSFGINQAVETNRDWSGTMKPTTDYAPALEYG 332
 DB 61 ELQVASTIVDSNGKVIAGLARGHSSNVSFGINQAVETNRDWSGTMKPTTDYAPALEYG 120
 OY 333 VYDSTATIVHDEPYNPGINTPYNMDRGYFGNITLQYALQOS 375

DB 121 VYDSTATIVHDEPYNPGINTPYNMDRGYFGNITLQYALQOS 163

RESULT 23
 O9E08 PRELIMINARY; PRT; 719 AA.
 AC O9E08:
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A.
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP 1465;
 RC MEDLINE-21432820; PubMed-11549185;
 RA Ferion A., Berche P.;
 RT "Alterations to penicillin-binding proteins 1A, 2B and 2X amongst
 penicillin-resistant clinical isolates of Streptococcus pneumoniae
 serotype 23F from the nasopharyngeal flora of children.";
 RT J. Med. Microbiol. 50:828-832(2001).
 RL EMBL; AF210747; AAF17257.1; -;
 DR EMBL; AF210747; AAF17257.1; -;
 DR InterPro; IPR001264; Transglycosyl.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00905; Transglycosyl; 1.
 DR Pfam; PF00905; Transpeptidase; 1.
 DR ProDom; PD001895; Transglycosyl; 1.
 SO SEQUENCE 719 AA; 79662 MW; A7B96EA15434A805 CRC64;

Query Match 21.3%; Score 142; DB 2; Length 719;
 Best Local Similarity 100.0%; Pred. No. 1.9e-135;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIDDKNKLADSGSRVNAQANDIPTDLKAIYSTEDHREFDNGIDIRILGAFLRN 60
 DB 54 KIDDKNKLADSGSRVNAQANDIPTDLKAIYSTEDHREFDNGIDIRILGAFLRN 113
 OY 61 LQSNLSLOGGSTLTQOLIKLTFSTSDQTSRKAQEWLAIOLEQATKQELITYYINK 120
 DB 114 LQSNLSLOGGSTLTQOLIKLTFSTSDQTSRKAQEWLAIOLEQATKQELITYYINK 173
 OY 121 VYMSNGNYGQTAQNYGKDL 142
 DB 174 VYMSNGNYGQTAQNYGKDL 195

RESULT 24
 O9E043 PRELIMINARY; PRT; 163 AA.
 AC O9E043:
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-950421;
 RC Overweg K., Bogaert D., Sluifster M., de Groot R., Hermans P.W.M.;
 RT "Molecular characterization of Streptococcus pneumoniae penicillin-
 resistance in the Netherlands.";
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ403977; CAC20959.1; -;

QY 560 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTESSSS 619
 DB 294 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTESSSS 353
 QY 620 SSDSSTSSQSSSTTPSTNNSTTNPNNNTQOOSNTTPOQONONPOPA 664
 DB 354 SSDSSTSSQSSSTTPSTNNSTTNPNNNTQOOSNTTPOQONONPOPA 398

RESULT 28

Q9R7M7 PRELIMINARY: PRT: 398 AA.

AC 09R7M7
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=#7/K133;
 RX MEDLINE=98409715; PubMed=9736547;
 RA Asahi Y., Ubukata K.;
 RT "Association of a Thr-371 substitution in a conserved amino acid motif of penicillin-binding protein 1A with penicillin resistance of Streptococcus pneumoniae."
 RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
 DR EMBL; AB006872; BAA32068.1; .
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 398
 SQ SEQUENCE 398 AA; 43638 MW; 551ADB567C3F79BE CRC64;

Query Match 15.8%; Score 105; DB 2; Length 398;
 Best Local Similarity 100.0%; Pred. No. 5.3e-98;
 Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTESSSS 619
 DB 294 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTESSSS 353
 QY 620 SSDSSTSSQSSSTTPSTNNSTTNPNNNTQOOSNTTPOQONONPOPA 664
 DB 354 SSDSSTSSQSSSTTPSTNNSTTNPNNNTQOOSNTTPOQONONPOPA 398

RESULT 29

Q87105 PRELIMINARY: PRT: 398 AA.

AC 087105
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=#9/217;
 RX MEDLINE=98409715; PubMed=9736547;
 RA Asahi Y., Ubukata K.;
 RT "Association of a Thr-371 substitution in a conserved amino acid motif of penicillin-binding protein 1A with penicillin resistance of Streptococcus pneumoniae."
 RT Streptococcus pneumoniae."

RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
 DR EMBL; AB006873; BAA32069.1; .
 DR InterPro: IPR001460; Transpeptidase.
 DR Pfam: PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 398
 SQ SEQUENCE 398 AA; 43727 MW; 3AA038C8B5582373 CRC64;

Query Match 15.8%; Score 105; DB 2; Length 398;
 Best Local Similarity 100.0%; Pred. No. 5.3e-98;
 Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 560 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTESSSS 619
 DB 294 VAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTESSSS 353
 QY 620 SSDSSTSSQSSSTTPSTNNSTTNPNNNTQOOSNTTPOQONONPOPA 664
 DB 354 SSDSSTSSQSSSTTPSTNNSTTNPNNNTQOOSNTTPOQONONPOPA 398

RESULT 30

Q54918 PRELIMINARY: PRT: 189 AA.

AC 054918
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE INTERNAL REGION OF THE PENICILLIN-BINDING PROTEIN 1A GENE (FRAGMENT).
 GN PBP 1A.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=85983 (PENICILLIN-RESISTANT STRAIN);
 RX MEDLINE=9401331; PubMed=8406829;
 RA Kell C.M., Jordens Z., Daniels M., Coffey T.J., Bates J., Paul J.,
 RA Cluks C., Spratt B.G.;
 RT "Molecular epidemiology of penicillin-resistant pneumococci isolated in Nairobi, Kenya."
 RT Infect. Immun. 61:4382-4391(1993).
 RL EMBL; Z21800; CAA79864.1; .
 FT NON_TER 1
 FT NON_TER 189
 SQ SEQUENCE 189 AA; 20740 MW; ADDDBCC22322A406D CRC64;

Query Match 14.7%; Score 98; DB 2; Length 189;
 Best Local Similarity 100.0%; Pred. No. 3.6e-91;
 Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 557 GLTYAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTES 616
 DB 92 GLTYAAKYRSMNTYLSGSGNPDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTES 151
 QY 617 SSSSDSSTSSQSSSTTPSTNNSTTNPNNNTQOOSNTT 654
 DB 152 SSSSDSSTSSQSSSTTPSTNNSTTNPNNNTQOOSNTT 189

RESULT 31

Q52737 PRELIMINARY: PRT: 310 AA.

AC 052737
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 GN PONA.

OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=65654;
RX MEDLINE=98287565; PubMed=9624469;
RA Smith A.M., Klugman K.P.;
RT "Alterations in PBP 1A essential for high-level penicillin resistance
in Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:1329-1333(1998).
DR EMBL: AF046231; AAC24696.1; -
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00905; Transpeptidase; 1.
FT NON_TER 1
FT 310 310
SQ SEQUENCE 310 AA; 34419 MW; 7C683053109F6DA CRC64;

Query Match 14.3%; Score 95; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 6.5e-68;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LMDIYNDEYVAYPPDELOVASTIVDSNGKVIQOLGARHOSNVSFGINQAVETNRDVG 316
|||||
DB 1 LMDIYNDEYVAYPPDELOVASTIVDSNGKVIQOLGARHOSNVSFGINQAVETNRDVG 60
QY 317 STMKPTIDYAPALEYGYDSTATIVHDEPYNPGT 351
|||||
DB 61 STMKPTIDYAPALEYGYDSTATIVHDEPYNPGT 95

RESULT 32
Q9R7M5 PRELIMINARY; PRT; 398 AA.
AC Q9R7M5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=#24/TJ25; #26/TJ29;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
DR EMBL: AB006879; BAA32075.1; -
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00905; Transpeptidase; 1.
FT NON_TER 1
FT 398 398
SQ SEQUENCE 398 AA; 43654 MW; D1E181574F2B58B8 CRC64;

Query Match 12.8%; Score 85; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.2e-77;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 VAYPDELOVASTIVDSNGKVIQOLGARHOSNVSFGINQAVETNRDVGSTMKPTIDYA 326
|||||
DB 1 VAYPDELOVASTIVDSNGKVIQOLGARHOSNVSFGINQAVETNRDVGSTMKPTIDYA 60
QY 327 PALEYGYDSTATIVHDEPYNPGT 351
|||||

DB 61 PALEYGYDSTATIVHDEPYNPGT 85
RESULT 33
ID 087108 PRELIMINARY; PRT; 398 AA.
AC 087108;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=#23/HSB21;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
Streptococcus pneumoniae.";
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
DR EMBL: AB006878; BAA32074.1; -
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00905; Transpeptidase; 1.
FT NON_TER 1
FT 398 398
SQ SEQUENCE 398 AA; 43675 MW; F513E2C548DC28AA CRC64;

Query Match 12.0%; Score 80; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.5e-72;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 HYSNMISSNTESDKKYGSSSEKMAAAYAFANGSTYKPMYIHKVPSDGSSEKFSNVG 466
|||||
DB 141 HYSNMISSNTESDKKYGSSSEKMAAAYAFANGSTYKPMYIHKVPSDGSSEKFSNVG 200
QY 467 TRAMKETAYMTDMKKTVL 486
|||||
DB 201 TRAMKETAYMTDMKKTVL 220

RESULT 34
O54950 PRELIMINARY; PRT; 719 AA.
AC O54950;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PONA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=2039;
RX MEDLINE=93010977; PubMed=1396576;
RA Martin C., Sibold C., Hakenbeck R.;
RT "Relatedness of penicillin-binding protein 1a genes from different
clones of penicillin-resistant Streptococcus pneumoniae isolated in
South Africa and Spain.";
RL EMBO J. 11:3831-3836(1992).
DR EMBL: X67870; CAA48070.1; -
DR InterPro: IPR001264; Transglycosyl.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.

DR ProDom: PD001895; Transglycosyl. 1.
FT NON_TER 719 719
SQ SEQUENCE 719 AA: 79874 MW: FACBFAB303D865D4 CRC64:

Query Match 11.9%; Score 79; DB 2; Length 719;
Best Local Similarity 100.0%; Pred. No. 2.8e-71;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 NSLOGSSTLTQOLIKLTFSTSTSDQFISRKAQEWLAIOLEQKAKFOELTYINKYVM 123
DB 117 NSLQGSSTLTQOLIKLTFSTSTSDQFISRKAQEWLAIOLEQKAKFOELTYINKYVM 176

QY 124 SNGNYGQTAQNYGKDL 142
DB 177 SNGNYGQTAQNYGKDL 195

RESULT 35
087107
ID 087107 PRELIMINARY: PRT: 398 AA.

DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).

GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=42/HA5.
RX MEDLINE=98409715; PubMed-9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae."
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).

DR EMBL; AB006877; BAA32073.1; -
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00905; Transpeptidase; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 398 AA: 43487 MW: CDF82616C5159991 CRC64:

Query Match 11.4%; Score 76; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.8e-68;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 589 RGEFFKNGASTWNSPAPQPPSTSSSSSDSTSSSSTPTTNNSTTNPNNNTQ 648
DB 323 RGEFFKNGASTWNSPAPQPPSTSSSSSDSTSSSSTPTTNNSTTNPNNNTQ 382

QY 649 OSNTTPDOONONPOPA 664
DB 383 OSNTTPDOONONPOPA 398

RESULT 36

09EM42
ID 09EM42 PRELIMINARY: PRT: 163 AA.

AC 09EM42;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).

GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.

OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=950473.

RA Overweg K., Bogaert D., Sluifster M., de Groot R., Hermans P.W.M.;
RT "Molecular characterization of Streptococcus pneumoniae penicillin-
resistance in the Netherlands."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ403979; CAC20961.1; -

FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 163 AA: 18350 MW: 02B96CAE2DB26ACF CRC64:

Query Match 9.0%; Score 60; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 1.6e-52;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 SNYPAYMDNYLKEVINQVEEETGYNLLTGMVYTNVDQEAQRHLDIYNTDEVYVPPD 272
DB 1 SNYPAYMDNYLKEVINQVEEETGYNLLTGMVYTNVDQEAQRHLDIYNTDEVYVPPD 60

RESULT 37
09R7M8
ID 09R7M8 PRELIMINARY: PRT: 398 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).

GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=43/R28, #6/KU26, #15/Z12, #16/Z13;
RX MEDLINE=98409715; PubMed-9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
RT Streptococcus pneumoniae."
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).

DR EMBL; AB006870; BAA32066.1; -
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00905; Transpeptidase; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 398 AA: 43583 MW: 0BD9B32C6B2D0577 CRC64:

Query Match 9.0%; Score 60; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 3.6e-52;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 605 SPAPQPPSTSSSSSDSTSSSSTPTTNNSTTNPNNNTQOSNTTPDOONONPOPA 664
DB 339 SPAPQPPSTSSSSSDSTSSSSTPTTNNSTTNPNNNTQOSNTTPDOONONPOPA 398

RESULT 38

09F2G4
ID 09F2G4 PRELIMINARY: PRT: 727 AA.

AC 09F2G4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).

GN PENIA.
OS Streptococcus mitis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Streptococcus.
 OX NCBI_TaxID=28037;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=209;
 RX MEDLINE=21301162; PubMed=11408226;
 RA Amoroso A., Demares D., Mollerach M., Gutkind G., Coyette J.;
 RT "All detectable high-molecular-mass penicillin-binding proteins are
 modified in a high-level beta-lactam-resistant clinical isolate of
 Streptococcus mitis";
 RL Antimicrob. Agents Chemother. 45:2075-2081(2001).
 DR EMBL; AJ295856; CAC08466.1; -;
 DR InterPro; IPR001264; Transglycosyl.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF009012; Transglycosyl; 1.
 DR Pfam; PF00905; Transpeptidase; 1.
 DR ProDom; PD001895; Transglycosyl; 1.
 FT NON_TER 727 727
 SQ SEQUENCE 727 AA; 80506 MW; AFF56454A4E30129 CRC64;

Query Match 8.1%; Score 54; DB 2; Length 727;
 Best Local Similarity 100.0%; Pred. No. 8.2e-46;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 285 NGKVTAGLGARHSSNSFGINQAVETNRDMSGTMKPTIDYAPALEYGYDSTA 338
 DB 339 NGKVTAGLGARHSSNSFGINQAVETNRDMSGTMKPTIDYAPALEYGYDSTA 392

RESULT 39
 ID 087104 PRELIMINARY; PRT; 398 AA.
 AC 087104;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=#5/H31, #8/22, #11/220, #12/221;
 RX MEDLINE=96409715; PubMed=9736547;
 RA Asahi Y., Ubukata K.;
 RT "Association of a Thr-371 substitution in a conserved amino acid motif
 of penicillin-binding protein 1A with penicillin resistance of
 Streptococcus pneumoniae";
 RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
 DR EMBL; AB006871; BAA32067.1; -;
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00905; Transpeptidase; 1.
 FT NON_TER 1 1
 FT NON_TER 398 398
 SQ SEQUENCE 398 AA; 43536 MW; 122DOB05F8EF117C CRC64;

Query Match 8.0%; Score 53; DB 2; Length 398;
 Best Local Similarity 100.0%; Pred. No. 4.9e-45;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 569 RGEFVFNGARSTWNSPAPQOPSTESSSSSDSTSSSTPSTNNSTTT 641
 DB 323 RGEFVFNGARSTWNSPAPQOPSTESSSSSDSTSSSTPSTNNSTTT 375

RESULT 40
 ID 09EM44 PRELIMINARY; PRT; 163 AA.
 AC 09EM44;

DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=950710;
 RA Overweg K., Bogaert D., Sluifder M., de Groot R., Hermans P.W.M.;
 RT "Molecular characterization of Streptococcus pneumoniae penicillin-
 resistance in the Netherlands";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ403976; CAC20958.1; -;
 FT NON_TER 1 1
 FT NON_TER 163 163
 SQ SEQUENCE 163 AA; 18358 MW; 3D50DD342D9A0900 CRC64;

Query Match 7.5%; Score 50; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 2.4e-42;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 213 SNYPAMDNYLKEVINQVEEETGYNLLTGMGVYTNVDDQAKHMDIYN 262
 DB 1 SNYPAMDNYLKEVINQVEEETGYNLLTGMGVYTNVDDQAKHMDIYN 50

RESULT 41
 ID 09EM41 PRELIMINARY; PRT; 163 AA.
 AC 09EM41;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
 GN PBPIA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=960027;
 RA Overweg K., Bogaert D., Sluifder M., de Groot R., Hermans P.W.M.;
 RT "Molecular characterization of Streptococcus pneumoniae penicillin-
 resistance in the Netherlands";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ403980; CAC20962.1; -;
 FT NON_TER 1 1
 FT NON_TER 163 163
 SQ SEQUENCE 163 AA; 18464 MW; 83007AD7ACB00B31 CRC64;

Query Match 6.2%; Score 41; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 3.6e-33;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 253 AOKHLMIDYNTDEYVAYPPDELQVASTIVDSNGKVIADLG 293
 DB 41 AOKHLMIDYNTDEYVAYPPDELQVASTIVDSNGKVIADLG 81

RESULT 42
 ID 093N72 PRELIMINARY; PRT; 310 AA.
 AC 093N72;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=45904;
RX MEDLINE=21393681; PubMed=11502545;
RA Smith A.M., Botha R.F., Koornhof H.J., Klugman K.P.;
RT "Emergence of a Pneumococcal Clone with Cephalosporin Resistance and
RT Penicillin Susceptibility."
RL Antimicrob. Agents Chemother. 45:2648-2650(2001).
DR EMBL; AF387163; AAK93962.1; -.
FT NON_TER 1 1
SQ SEQUENCE 310 AA; 34249 MW; 9756459BCFA2EF CRC64;

Query Match 6.2%; Score 41; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 6.5e-33;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 257 LMDIYNTDEVAYPDDELQVASTIVDSNGKYIAQIGARHQ 297
Db 1 LMDIYNTDEVAYPDDELQVASTIVDSNGKYIAQIGARHQ 41

RESULT 43
O70037 PRELIMINARY; PRT; 527 AA.
AC O70037;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus mitis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=28037;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6;
RX MEDLINE=98196728; PubMed=9537382;
RA Hakenbeck R., Koenig A., Kern I., van der Linden M., Keck W.,
RA Billot-Klein D., Legrand R., Schoot B., Gutmann L.;
RT "Acquisition of five high-Mr penicillin-binding protein variants
RT during transfer of high-level beta-lactam resistance from
RT Streptococcus mitis to Streptococcus pneumoniae."
RL J. Bacteriol. 180:1831-1840(1998).
DR EMBL; AJ002290; CAA05301.1; -.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 1 1
FT NON_TER 527 527
SQ SEQUENCE 527 AA; 58859 MW; 8952B38713827021 CRC64;

Query Match 5.7%; Score 38; DB 2; Length 527;
Best Local Similarity 100.0%; Pred. No. 1.2e-29;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 NSLGGSTLTQOLIKITFTSTSDOTISRAQEAAMLA 101
Db 67 NSLGGSTLTQOLIKITFTSTSDOTISRAQEAAMLA 104

RESULT 44
O9F2G7

ID O9F2G7 PRELIMINARY; PRT; 728 AA.
AC O9F2G7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PENIA.
OS Streptococcus mitis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=28037;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=127R;
RX MEDLINE=21301162; PubMed=11408226;
RA Amoroso A., Demares D., Moellerach M., Gutkind G., Coyette J.;
RT "All detectable high-molecular-mass penicillin-binding proteins are
RT modified in a high-level beta-lactam-resistant clinical isolate of
RT Streptococcus mitis."
RL Antimicrob. Agents Chemother. 45:2075-2081(2001).
DR EMBL; AJ295852; CAC08463.1; -.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
FT NON_TER 728 728
SQ SEQUENCE 728 AA; 80854 MW; 37E110286600987 CRC64;

Query Match 4.8%; Score 32; DB 2; Length 728;
Best Local Similarity 100.0%; Pred. No. 2.1e-23;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 103 QLEQKATKOEILTYINKYMSNGYGMOTAA 134
Db 157 QLEQKATKOEILTYINKYMSNGYGMOTAA 188

RESULT 45
O9EW45 PRELIMINARY; PRT; 163 AA.
AC O9EW45;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=950225;
RA Overweg K., Bogaert D., Sluifjter M., de Groot R., Hermans P.W.M.;
RT "Molecular characterization of Streptococcus pneumoniae penicillin-
RT resistance in the Netherlands."
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF403974; CAC20956.1; -.
FT NON_TER 1 1
FT NON_TER 163 163
SQ SEQUENCE 163 AA; 18357 MW; F6DED7F42D80D3CB CRC64;

Query Match 4.5%; Score 30; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 5.7e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 233 ETGYNLLTGMVYTNVDOEAKHMDIYN 262
Db 21 ETGYNLLTGMVYTNVDOEAKHMDIYN 50

RESULT 46
087106 PRELIMINARY; PRT; 400 AA.
ID 087106;
AC 087106;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN#14/242;
RX MEDLINE=98409715; PubMed=9736547;
RA Asahi Y., Ubukata K.;
RT "Association of a Thr-371 substitution in a conserved amino acid motif
of penicillin-binding protein 1A with penicillin resistance of
Streptococcus pneumoniae."
RL Antimicrob. Agents Chemother. 42:2267-2273(1998).
DR EMBL; AB006875; BAA32071.1; -
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00905; Transpeptidase; 1.
FT NON_TER 1 1
FT NON_TER 400 400
SO SEQUENCE 400 AA; 43773 MW; 0B8965011453D226 CRC64;

Query Match 4.1%; Score 27; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 1,5e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 267 VAYPDELQVASTIVDSNGKVIQOLG 293
Db 1 VAYPDELQVASTIVDSNGKVIQOLG 27
|||||

RESULT 47
099YL1 PRELIMINARY; PRT; 721 AA.
ID 099YL1;
AC 099YL1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PURATIVE PENICILLIN-BINDING PROTEIN 1A.
GN PBPIA OR SPY1649.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., Moshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Szate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006596; AAK34416.1; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Complete proteome.
SO SEQUENCE 721 AA; 80207 MW; 7E9350705EF4C6D0 CRC64;

Query Match 3.5%; Score 23; DB 16; Length 721;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 419 SPKKGASSEKMAAAYAFANGG 441
Db 475 SDKKYGASSEKMAAAYAFANGG 497
|||||

RESULT 48
Q9EW40 PRELIMINARY; PRT; 163 AA.
ID Q9EW40;
AC Q9EW40;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PENICILLIN BINDING PROTEIN 1A (FRAGMENT).
GN PBPIA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-960035;
RA Overweg K., Bogaert D., Sluifder M., de Groot R., Hermans P.W.M.;
RT "Molecular characterization of Streptococcus pneumoniae penicillin-
resistance in the Netherlands."
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ403981; CAC20963.1; -
FT NON_TER 1 1
FT NON_TER 163 163
SO SEQUENCE 163 AA; 18347 MW; BC800A043C6E1FCF CRC64;

Query Match 3.3%; Score 22; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 8e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 241 TGMVYTNVDQAKKHLMDIYN 262
Db 29 TGMVYTNVDQAKKHLMDIYN 50
|||||

RESULT 49
Q54919 PRELIMINARY; PRT; 139 AA.
ID Q54919;
AC Q54919;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INTERNAL REGION OF THE PENICILLIN-BINDING PROTEIN 1A GENE
(FRAGMENT).
GN PBP 1A.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-100511 (PENICILLIN-RESISTANT STRAIN);
RX MEDLINE=94011331; PubMed=8406829;
RA Kell C.M., Jordens Z., Daniels M., Coffey T.J., Bates J., Paul J.,
RA Gilks C., Spratt B.G.;
RT "Molecular epidemiology of penicillin-resistant pneumococci isolated
in Nairobi, Kenya."
RL Infect. Immun. 61:4382-4391(1993).
DR EMBL; Z21809; CAA79874.1; -
FT NON_TER 1 1
FT NON_TER 139 139
SO SEQUENCE 139 AA; 15250 MW; 40ED8ED03593D446 CRC64;

Query Match 3.0%; Score 20; DB 2; Length 139;
 Best Local Similarity 100.0%; Pred. No. 7.6e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 533 GYRKYSMAVMTGYSNRLTP 552
 ||||||||||||||||
 Db 16 GYRKYSMAVMTGYSNRLTP 35

RESULT 50

052744 PRELIMINARY; PRT; 310 AA.
 ID 052744
 AC 052744;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PENICILLIN-BINDING PROTEIN 1A (FRAGMENT).
 GN PONA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=63509, M11;
 RX MEDLINE=98287565; PubMed=9624469;
 RA Smith A.M., Klugman K.P.;
 RT "Alterations in PBP 1A essential for high-level penicillin resistance
 in Streptococcus pneumoniae."
 RL Antimicrob. Agents Chemother. 42:1329-1333(1998).
 DR EMBL; AF046238; AAC24703.1; -;
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00905; Transpeptidase; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 310 AA; 34354 MW; F8FC8AC00E7BE5FF CRC64;

Query Match 3.0%; Score 20; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 1.6e-11;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 MKPTDYAPALEGYDSTA 338
 ||||||||||||||||
 Db 63 MKPTDYAPALEGYDSTA 82

Search completed: June 13, 2002, 08:50:23
 Job time: 238 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:42:44 ; Search time 14.67 Seconds
(without alignments)
1108.892 Million cell updates/sec

Title: US-08-961-083-2
Perfect score: 666
Sequence: 1 KIYDNKNQLIADLGSERRRN.....TOOSNTTPOOONPQAPQ 666

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 10

Total number of hits satisfying chosen parameters: 8

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database :

Issued_Patents_AA:*
1: /cgn2-6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2-6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2-6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2-6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2-6/ptodata/2/1aa/PTUTS.COMB.pep:*
6: /cgn2-6/ptodata/2/1aa/Backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	666	100.0	666	4	US-08-961-083-2
2	330	49.5	682	3	US-08-481-435-6
3	245	36.8	320	2	US-08-245-511-4
4	245	36.8	320	2	US-08-600-993A-4
5	77	11.6	77	2	US-08-245-511-24
6	77	11.6	77	2	US-08-600-993A-24
7	10	1.5	110	4	US-08-961-083-102
8	10	1.5	1704	4	US-08-485-355B-40

ALIGNMENTS

RESULT 1
US-08-961-083-2
; Sequence 2, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA

ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: P834072
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-2

Query Match 100.0%; Score 666; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KIYDNKNQLIADLGSERRRNANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLLN	60
DB	1	KIYDNKNQLIADLGSERRRNANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLLN	60
QY	61	IQSNLSQGGSTLTQOLIKITFYSTSDOTISRKQAEMLAIQLEKAKROELTYIYK	120
DB	61	IQSNLSQGGSTLTQOLIKITFYSTSDOTISRKQAEMLAIQLEKAKROELTYIYK	120
QY	121	YVMSNGNYGMOTAAQNYVYKDLNNLSLPOLALLAGPQAPNOYDPYSHPEAODRNLV	180
DB	121	YVMSNGNYGMOTAAQNYVYKDLNNLSLPOLALLAGPQAPNOYDPYSHPEAODRNLV	180
QY	181	SEMKNQYISADQYERAVNTPTDGLQSLKSASNPAYMDNYLKEVINOVEERTGYNLT	240
DB	181	SEMKNQYISADQYERAVNTPTDGLQSLKSASNPAYMDNYLKEVINOVEERTGYNLT	240
QY	241	TGMQVYTNDOAKHMDIYNTDEVAVTPDELOVASTIVVSNKVIYAOIGARHOSN	300
DB	241	TGMQVYTNDOAKHMDIYNTDEVAVTPDELOVASTIVVSNKVIYAOIGARHOSN	300
QY	301	VSEGINQAVETNRDNGSTMKPTIDYAPALEYGVYDSTATIVHDEPNYPTGNTPYNMNR	360
DB	301	VSEGINQAVETNRDNGSTMKPTIDYAPALEYGVYDSTATIVHDEPNYPTGNTPYNMNR	360
QY	361	GYFGNTTQYALQOQRNPAVEETLNVKGLNRAKTEFLNGIDYPSITHYSNAISSNTTESD	420
DB	361	GYFGNTTQYALQOQRNPAVEETLNVKGLNRAKTEFLNGIDYPSITHYSNAISSNTTESD	420
QY	421	KRYGASSEKMAAAYAFAGGTYRPMYTHKVFSDGSEKESNNGTRAMKETTAAMD	480
DB	421	KRYGASSEKMAAAYAFAGGTYRPMYTHKVFSDGSEKESNNGTRAMKETTAAMD	480
QY	481	NKTYLTYGTGRNAYLAMPQAGKTGTSNYTDEEINHIKTQFVAPDELFGYTRKYSM	540
DB	481	NKTYLTYGTGRNAYLAMPQAGKTGTSNYTDEEINHIKTQFVAPDELFGYTRKYSM	540
QY	541	AVMTGYSNRLPLVNGILTVAAKVVYRSMNTYLSGSPNEDMNIPEGLYRNGEFVRKNGAR	600
DB	541	AVMTGYSNRLPLVNGILTVAAKVVYRSMNTYLSGSPNEDMNIPEGLYRNGEFVRKNGAR	600

QY	601	STWSPAPPOOPSTESSSSSSDSTSQSSSTPTNNSTTNNNNNQSSNTDPDOON	666
Db	601	STWSPAPPOOPSTESSSSSSDSTSQSSSTPTNNSTTNNNNNQSSNTDPDOON	666
QY	661	POPAP	666
Db	661	POPAP	666

RESULT 2

```

US-08-481-435-6
Sequence 6, Application US/08481435.
Patent No. 6027906
GENERAL INFORMATION:
APPLICANT: Balganesht, Tanjore S
APPLICANT: Town, Christine
TITLE OF INVENTION: No. 6027906el Polypeptides
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: White & Case
STREET: 1155 Avenue of the Americas
City: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,435
FILING DATE: 10-JUL-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IN 580/MAS/94
FILING DATE: 01-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION DATA: SE 9404072-2
APPLICATION NUMBER: 24-NOV-1994
FILING DATE: 1994
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 682 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-435-6

```

Query Match	49.5%	Score 330	DB 3	Length 682
-------------	-------	-----------	------	------------

[illegible]

Qy	252	EAGHLMNDIYTDDEYVYPPDELOVASTIYDVNSGKIYALQAGRHOSNVSFGINAVET	31.1
Db	268	EAGHLMNDIYTDDEYVYPPDELOVASTIYDVNSGKIYALQAGRHOSNVSFGINAVET	32.7
Qy	312	NRDMGSTMKPITDYAPALEGVYDSTATIYHDEPYNTPGINTPYVMMDRGYFGNITLYQYA	37.1
Db	328	NRDMGSTMKPITDYAPALEGVYDSTATIYHDEPYNPGINTPYVMMDRGYFGNITLYQYA	38.7
Qy	372	LQOSRNPVAVETLKKVGNRAKRTPLNGLGIDYPSIYHSNMAISSNTWESDCKKYGASSEKMA	43.1
Db	388	LQOSRNPVAVETLKKVGNRAKRTPLNGLGIDYPSIYHSNMAISSNTWESDCKKYGASSEKMA	44.7
Qy	432	AAATAPFANGGTYYKPYNIHKYVESDGSSEKESNNGTAPAKMETTAYMMTDMKITYLYTGTG	49.1
Db	448	AAATAPFANGGTYYKPYNIHKYVESDGSSEKESNNGTAPAKMETTAYMMTDMKITYLYTGTG	50.7
Qy	492	RNATLAWLPOAGKGTGSNTYDDEEIEHNIIKTSOFVAPDELPAGYTRKYSMAVMTGYSNRLT	55.1
Db	508	RNATLAWLPOAGKGTGSNTYDDEEIEHNIIKTSOFVAPDELPAGYTRKYSMAVMTGYSNRLT	56.7
Qy	552	PLVNGGLTVAAKYRVKSMATYLSBGSNPNEDMNIIPGATLRNPEFVFKNCARSTW	60.3
Db	568	PLVNGGLTVAAKYRVKSMATYLSBGSNPNEDMNIIPGATLRNPEFVFKNCARSTW	61.9

RESULT 3

```

US-08-245-511-4
: Sequence 4, Application US/08245511
: Patent No. 5928900
:
: GENERAL INFORMATION:
:
: APPLICANT: Masure, H Robert
: APPLICANT: Pearce, Barbara J
: APPLICANT: Tucumanen, Elaine
: TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS
: TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERE
:
: NUMBER OF SEQUENCES: 58
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Klauber & Jackson
: STREET: 411 Hackensack Avenue
: CITY: Hackensack
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07601
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/245,511
: FILING DATE: 18-MAY-1994
: CLASSIFICATION: 424
:
: PRIOR APPLICATION DATA:
:
: APPLICATION NUMBER: US 08/116,541
: FILING DATE: 01-SEP-1994
:
: ATTORNEY/AGENT INFORMATION:
:
: NAME: Jackson Esq., David A.
: REGISTRATION NUMBER: 26,742
: REFERENCE/DOCKET NUMBER: 600-1-069 CIP
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: 201 487-5800
: TELEFAX: 201 343-1684
:
: TELEX: 133521
:
: INFORMATION FOR SEQ ID NO: 4:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 320 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
: US-08-245-511-4

```

Query Match 36.8%; Score 245; DB 2; Length 320;

Best Local Similarity 100.0%; Pred. No. 1.3e-241;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 72 LKQQLIKLTYFSTSDQITSRKAOEAWLAIOLEOKATKOEILTYINKVYNSNGNYGMQ 131
Db 76 LKQQLIKLTYFSTSDQITSRKAOEAWLAIOLEOKATKOEILTYINKVYNSNGNYGMQ 135
OY 132 TAAQNYGGDLNLSLPQALLAGMPQAPNOYDPYSHPEAADRRNLVISEMKNQGYISA 191
Db 136 TAAQNYGGDLNLSLPQALLAGMPQAPNOYDPYSHPEAADRRNLVISEMKNQGYISA 195
OY 192 EOYKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINQVEEETGYNLTTGMDVYTNVQ 251
Db 196 EOYKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINQVEEETGYNLTTGMDVYTNVQ 255
OY 252 EAQKHLMDIYNDEYVAYPDDELQVASTIVDVSNGKVIQOLGARHSSNVSGINQAVET 311
Db 256 EAQKHLMDIYNDEYVAYPDDELQVASTIVDVSNGKVIQOLGARHSSNVSGINQAVET 315
OY 312 NRDMG 316
Db 316 NRDMG 320
```

RESULT 4

US-08-600-993A-4
; Sequence 4, Application US/08600993A
; Patent No. 5981229

GENERAL INFORMATION:

APPLICANT: Masure, H Robert
APPLICANT: Pearce, Barbara J
APPLICANT: Tuomanen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,993A
FILING DATE: 1-MAR-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-600-993A-4

Query Match 36.8%; Score 245; DB 2; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.3e-241;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 72 LKQQLIKLTYFSTSDQITSRKAOEAWLAIOLEOKATKOEILTYINKVYNSNGNYGMQ 131
Db 76 LKQQLIKLTYFSTSDQITSRKAOEAWLAIOLEOKATKOEILTYINKVYNSNGNYGMQ 135
OY 132 TAAQNYGGDLNLSLPQALLAGMPQAPNOYDPYSHPEAADRRNLVISEMKNQGYISA 191
Db 136 TAAQNYGGDLNLSLPQALLAGMPQAPNOYDPYSHPEAADRRNLVISEMKNQGYISA 195
OY 192 EOYKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINQVEEETGYNLTTGMDVYTNVQ 251
Db 196 EOYKAVNTPITDGLQSLKSASNYPAYMDNYLKEVINQVEEETGYNLTTGMDVYTNVQ 255
OY 252 EAQKHLMDIYNDEYVAYPDDELQVASTIVDVSNGKVIQOLGARHSSNVSGINQAVET 311
Db 256 EAQKHLMDIYNDEYVAYPDDELQVASTIVDVSNGKVIQOLGARHSSNVSGINQAVET 315
OY 312 NRDMG 316
Db 316 NRDMG 320
```

RESULT 5

US-08-245-511-24
; Sequence 24, Application US/08245511
; Patent No. 5928900

GENERAL INFORMATION:

APPLICANT: Masure, H Robert
APPLICANT: Pearce, Barbara J
APPLICANT: Tuomanen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
TOPOLOGY: unknown

MOLECULE TYPE: peptide
HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: R6
IMMEDIATE SOURCE:
CLONE: SPRU42
US-08-245-511-24

Query Match 11.6%; Score 77; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 8.1e-71;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 TTGMDVYTNVDOEAQKHLMDIYNTDEVVAYPPDELOVASTIVVSNCKVIAQLGARHSS 299
DB 1 TTGMDVYTNVDOEAQKHLMDIYNTDEVVAYPPDELOVASTIVVSNCKVIAQLGARHSS 60

QY 300 NVSFGINQAVETNRDMG 316
DB 61 NVSFGINQAVETNRDMG 77

RESULT 6

US-08-600-993A-24
Sequence 24, Application US/08600993A
Patent No. 5981229
GENERAL INFORMATION:
APPLICANT: Measure, H Robert
APPLICANT: Pearce, Barbara J
APPLICANT: Tuomanen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,993A
FILING DATE: 1-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,511
FILING DATE: 18-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/116,541
FILING DATE: 01-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
MOLECULE TYPE: unknown
TOPOLOGY: unknown
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: R6
IMMEDIATE SOURCE:
CLONE: SPRU42
US-08-600-993A-24

Query Match 11.6%; Score 77; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 8.1e-71;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 TTGMDVYTNVDOEAQKHLMDIYNTDEVVAYPPDELOVASTIVVSNCKVIAQLGARHSS 299
DB 1 TTGMDVYTNVDOEAQKHLMDIYNTDEVVAYPPDELOVASTIVVSNCKVIAQLGARHSS 60

QY 300 NVSFGINQAVETNRDMG 316
DB 61 NVSFGINQAVETNRDMG 77

RESULT 7

US-08-961-083-102
Sequence 102, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-083-102

Query Match 1.5%; Score 10; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 18 SSSSSSDSST 27

RESULT 8
US-08-485-355B-40
; Sequence 40, Application US/08485355B
; Patent No. 6177075
; GENERAL INFORMATION:
; APPLICANT: Christlian, P. D., Gordon, K. H.J., Hanzlik, T. N.
; TITLE OF INVENTION: Insect Viruses and Their Uses in
; Protecting Plants
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr Hohbach Test Albritton & Herbert LLP
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,355B
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,522
; FILING DATE: 12-MAY-1995
; APPLICATION NUMBER: US 08/089,372
; FILING DATE: 08-JUL-1993
; APPLICATION NUMBER: AU PL4081/92
; FILING DATE: 14-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-58631-2/RT/DSS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1704 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
; US-08-485-355B-40

Query Match 1.5%; Score 10; DB 4; Length 1704;
Best Local Similarity 100.0%; Pred.No. 0.38;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 617 SSSSSDSSTS 626
|||||
DB 1452 SSSSSDSSTS 1461

Search completed: June 13, 2002, 08:46:41
Job time: 237 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:44:44 ; Search time 113.7 Seconds

(without alignments)
2061.733 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666

Sequence: 1 KIDYKNGQIADIGSERRVN.....TQSSNTTPDQGNQNPQAPQ 666

Scoring table: OLIGO

Gapop 60.0 , Gapect 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 10

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCrUS.COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081.COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082.COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083.COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084.COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085.COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087.COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088.COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089.COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090.COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091.COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092.COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093.COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094.COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095.COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096.COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097.COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098.COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099.COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US100.COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US101.COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US60.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	666	100.0	666	19	US-09-536-784-2
2	666	100.0	666	21	US-09-765-271-2
3	666	100.0	666	21	US-09-765-272-2
4	565	84.8	719	19	US-09-583-110-3973
5	565	84.8	721	15	US-09-107-433-3705
6	330	49.5	719	1	PCT-US02-03987-13423
7	330	49.5	719	22	US-09-815-242-13423

8	330	49.5	719	24	US-10-072-851-13423	Sequence 13423, A
9	245	36.8	777	5	US-08-116-541-4	Sequence 4, Appl1
10	77	11.6	823	15	US-08-116-541-24	Sequence 24, Appl1
11	18	2.7	823	15	US-09-107-532-5667	Sequence 5667, Ap
12	18	2.7	823	15	US-09-107-532A-5667	Sequence 5667, Ap
13	13	2.1	771	20	US-09-634-238-229	Sequence 229, Ap
14	14	2.1	771	20	US-09-634-238-262	Sequence 262, Ap
15	14	2.1	778	1	PCT-US02-03987-10728	Sequence 10728, A
16	14	2.1	778	22	US-09-815-242-10728	Sequence 10728, A
17	14	2.1	778	24	US-10-072-851-10728	Sequence 10728, A
18	14	2.1	789	15	US-09-134-000-4939	Sequence 4939, Ap
19	19	1.7	266	13	PCT-US97-14436-535	Sequence 535, App
20	11	1.7	266	13	US-08-911-503-535	Sequence 535, App
21	11	1.7	266	13	US-08-911-503A-535	Sequence 535, App
22	11	1.7	314	18	US-09-417-507-35782	Sequence 35782, A
23	11	1.7	820	19	US-09-583-110-4219	Sequence 4219, Ap
24	11	1.7	821	17	US-09-377-465-2	Sequence 2, Appl1
25	11	1.7	821	17	US-09-377-465A-2	Sequence 2, Appl1
26	11	1.7	834	15	US-09-107-433-4760	Sequence 4760, Ap
27	10	1.5	76	20	US-09-621-976-7706	Sequence 7706, Ap
28	10	1.5	76	22	US-09-834-366-24930	Sequence 24930, A
29	10	1.5	76	22	US-09-834-366-24930	Sequence 24930, A
30	10	1.5	76	26	US-06-147-499-7706	Sequence 7706, Ap
31	10	1.5	76	26	US-06-197-873-24929	Sequence 24929, A
32	10	1.5	76	26	US-06-197-873-24929	Sequence 24929, A
33	10	1.5	110	21	US-09-536-784-102	Sequence 102, App
34	10	1.5	110	21	US-09-536-784-102	Sequence 102, App
35	10	1.5	110	21	US-09-765-271-102	Sequence 102, App
36	10	1.5	119	20	US-09-614-150-34311	Sequence 34311, A
37	10	1.5	119	20	US-06-191-637-33895	Sequence 33895, A
38	10	1.5	130	20	US-09-621-976-5866	Sequence 5866, Ap
39	10	1.5	130	26	US-06-147-499-5866	Sequence 5866, Ap
40	10	1.5	194	23	US-09-950-083-3079	Sequence 3079, Ap
41	10	1.5	263	1	PCT-US01-27604-504	Sequence 504, App
42	10	1.5	263	1	PCT-US01-27604-504	Sequence 504, App
43	10	1.5	263	21	US-09-790-893-3	Sequence 6, Appl1
44	10	1.5	284	16	US-09-056-019-3	Sequence 3, Appl1
45	10	1.5	284	16	US-09-287-788-3	Sequence 3, Appl1
46	10	1.5	406	14	US-09-056-019-1	Sequence 1, Appl1
47	10	1.5	406	14	US-09-287-788-1	Sequence 1, Appl1
48	10	1.5	428	16	US-09-056-019-24	Sequence 24, Appl1
49	10	1.5	428	16	US-09-287-788-24	Sequence 24, Appl1
50	10	1.5	428	20	US-09-605-7038-1256	Sequence 1256, Ap
51	10	1.5	428	20	US-09-605-7038-1258	Sequence 1258, Ap
52	10	1.5	438	21	US-09-738-626-6703	Sequence 6703, Ap
53	10	1.5	446	16	US-09-286-9818-6	Sequence 6, Appl1
54	10	1.5	446	16	US-09-286-9818-9	Sequence 9, Appl1
55	10	1.5	459	14	US-09-056-019-39	Sequence 39, Appl1
56	10	1.5	459	16	US-09-287-788-39	Sequence 39, Appl1
57	10	1.5	460	14	US-09-056-019-38	Sequence 38, Appl1
58	10	1.5	460	16	US-09-287-788-38	Sequence 38, Appl1
59	10	1.5	462	24	US-10-091-007-92	Sequence 92, Appl1
60	10	1.5	511	16	US-09-298-523B-13	Sequence 13, Appl1
61	10	1.5	511	16	US-09-298-523C-13	Sequence 13, Appl1
62	10	1.5	511	21	US-09-748-875-13	Sequence 13, Appl1
63	10	1.5	513	16	US-09-298-523B-12	Sequence 12, Appl1
64	10	1.5	513	16	US-09-298-523C-12	Sequence 12, Appl1
65	10	1.5	513	21	US-09-748-875-12	Sequence 12, Appl1
66	10	1.5	655	14	US-09-056-019-2	Sequence 2, Appl1
67	10	1.5	655	16	US-09-287-788-2	Sequence 2, Appl1
68	10	1.5	693	21	US-09-769-787-185	Sequence 185, App
69	10	1.5	698	22	US-09-739-449-11825	Sequence 11825, A
70	10	1.5	698	22	US-09-803-110-11825	Sequence 11825, A
71	10	1.5	757	21	US-09-739-449-11939	Sequence 11939, A
72	10	1.5	757	22	US-09-803-110-11939	Sequence 11939, A
73	10	1.5	774	1	PCT-US02-03987-12046	Sequence 12046, A
74	10	1.5	774	19	US-09-543-681A-5851	Sequence 5851, Ap
75	10	1.5	774	22	US-09-815-242-12046	Sequence 12046, A
76	10	1.5	774	22	US-10-072-851-12046	Sequence 12046, A
77	10	1.5	777	16	US-09-252-991A-23697	Sequence 23697, A
78	10	1.5	781	1	PCT-US02-03987-11277	Sequence 11277, A
79	10	1.5	781	22	US-09-815-242-11277	Sequence 11277, A
80	10	1.5	781	24	US-10-072-851-11277	Sequence 11277, A

81	10	1.5	814	17	US-09-328-352-4373	Sequence 4373, Ap
82	10	1.5	825	22	US-09-897-516-7108	Sequence 7108, Ap
83	10	1.5	825	26	US-60-215-161-7108	Sequence 7108, Ap
84	10	1.5	1704	4	US-08-089-372-40	Sequence 40, Appl
85	10	1.5	1704	8	US-08-440-522-40	Sequence 40, Appl
86	10	1.5	1704	8	US-08-440-522a-40	Sequence 40, Appl
87	10	1.5	1704	8	US-08-485-355a-40	Sequence 40, Appl
88	10	1.5	1704	16	US-09-234-238-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1

US-09-536-784-2

Sequence 2, Application US/09536784

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSER: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 666 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-536-784-2

Query Match 100.0%; Score 666; DB 19; Length 666;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KIYDNKNOLIADGSRERVAQANDIPTDLVKAIVSIEDHRFPDHGIDTIRILGAFLRN	60
DB	1	KIYDNKNOLIADGSRERVAQANDIPTDLVKAIVSIEDHRFPDHGIDTIRILGAFLRN	60
QY	61	LOSNSLQGGSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIOLEQKATKQEIILTYINK	120
DB	61	LOSNSLQGGSTLTQOLIKLTYFSTSDQTSIRKQAEAMLAIOLEQKATKQEIILTYINK	120
QY	121	VYMSNGNYGMQTAQNYVGGDLNNLSLPOLALLAGMPQAPNOVDPTSHPEAADRRRLV	180
DB	121	VYMSNGNYGMQTAQNYVGGDLNNLSLPOLALLAGMPQAPNOVDPTSHPEAADRRRLV	180

QY	181	SEKKNQGYISAEQYERKAVNPTITDGLQSLKSNAPYAMDNLYKEVINQVEEFTGNLLT	240
DB	181	SEKKNQGYISAEQYERKAVNPTITDGLQSLKSNAPYAMDNLYKEVINQVEEFTGNLLT	240
QY	241	TGMDVYTNVQEQAKHLMIDYNTDEYVAYPPDELOVASTIVDSNGKVIAOLGARHOSSN	300
DB	241	TGMDVYTNVQEQAKHLMIDYNTDEYVAYPPDELOVASTIVDSNGKVIAOLGARHOSSN	300
QY	301	VSFGINQAVEETNRDQSTMKPTIDYAPALEGYVDSTATIVHDEPYNYPCTNTPVYMWDR	360
DB	301	VSFGINQAVEETNRDQSTMKPTIDYAPALEGYVDSTATIVHDEPYNYPCTNTPVYMWDR	360
QY	361	GFEGNITLQYALQOSRNVPAVELNKKVLRARFTFLNGLIDYPSIHYNSAISNTTESD	420
DB	361	GFEGNITLQYALQOSRNVPAVELNKKVLRARFTFLNGLIDYPSIHYNSAISNTTESD	420
QY	421	KKYGASSEKMAAAYAFANGTYKKPYIHKYVPSDSSEKFEPSNVGTRAKKETTAYMTD	480
DB	421	KKYGASSEKMAAAYAFANGTYKKPYIHKYVPSDSSEKFEPSNVGTRAKKETTAYMTD	480
QY	481	MMKTVLTYGTGRNAYLAMLPOAGKTGTSNTDEIEIENHIKTSQFVAPDELFACTYTRYSM	540
DB	481	MMKTVLTYGTGRNAYLAMLPOAGKTGTSNTDEIEIENHIKTSQFVAPDELFACTYTRYSM	540
QY	541	AVWTGYSNRLTPLYNGNLVAAKVRSMYTLSEGSNPEDMNIPEGLYRNGEEVFKNGAR	600
DB	541	AVWTGYSNRLTPLYNGNLVAAKVRSMYTLSEGSNPEDMNIPEGLYRNGEEVFKNGAR	600
QY	601	STWNSPAPQPPSTESSSSSDSSTTSQSSSTTSTNNSTTTNNNTQCSNTTTPDOONN	660
DB	601	STWNSPAPQPPSTESSSSSDSSTTSQSSSTTSTNNSTTTNNNTQCSNTTTPDOONN	660
QY	661	POPAOP 666	
DB	661	POPAOP 666	

RESULT 2

US-09-765-271-2

Sequence 2, Application US/09765271

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSER: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,271

FILING DATE: 22-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/536,784

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB340P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-271-2

Query Match 100.0%; Score 666; DB 21; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYDKNKNLIADLGSERRNAOANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLELN 60
DB 1 KYDKNKNLIADLGSERRNAOANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLELN 60
QY 61 LOSNSLGGSTLTQOLIKLTFSTSTSDQTSIRKAOEAMLAIOLEKATKOEILTYIYNK 120
DB 61 LOSNSLGGSTLTQOLIKLTFSTSTSDQTSIRKAOEAMLAIOLEKATKOEILTYIYNK 120
QY 121 VYMSNGNTGMOTAONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPDEAADRRLVL 180
DB 121 VYMSNGNTGMOTAONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPDEAADRRLVL 180
QY 181 SEMKNOGYISAQOYERKAVNPTITDGLQSLKASNPAYMDNLYKEVINOVEETGYNLT 240
DB 181 SEMKNOGYISAQOYERKAVNPTITDGLQSLKASNPAYMDNLYKEVINOVEETGYNLT 240
QY 241 TGMDEVYTNDOAOKHLMIDYNTDEVVAYPPDELOVASTIVDSNGKVAIOLGARHOSN 300
DB 241 TGMDEVYTNDOAOKHLMIDYNTDEVVAYPPDELOVASTIVDSNGKVAIOLGARHOSN 300
QY 301 VSFGINQAVETNRDMSGTMRKPTTDYAPALEYGYVDSTATIVHDEPNYPGTNPVYNMNR 360
DB 301 VSFGINQAVETNRDMSGTMRKPTTDYAPALEYGYVDSTATIVHDEPNYPGTNPVYNMNR 360
QY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFNLGIDYPSIHYSNAISSNTTESD 420
DB 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFNLGIDYPSIHYSNAISSNTTESD 420
QY 421 KKYGASSEKMAAAYAFANGGTYYKPMYTHKVFSDGSEKESNVTGRAMKETTAAYMMTD 480
DB 421 KKYGASSEKMAAAYAFANGGTYYKPMYTHKVFSDGSEKESNVTGRAMKETTAAYMMTD 480
QY 481 MKKTVLTGTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
DB 481 MKKTVLTGTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
QY 541 AVMTGYSNRLPPLVGNGLTVAAKYVRSMWTYLSSEGSNPEDMNIPEGIVRNGSEFVRKNGAR 600
DB 541 AVMTGYSNRLPPLVGNGLTVAAKYVRSMWTYLSSEGSNPEDMNIPEGIVRNGSEFVRKNGAR 600
QY 601 STWNSPAPQOPSTESSSSSDSTSSQSSSTPTNNNSTTTPNNNTQOOSNTTPOQONN 660
DB 601 STWNSPAPQOPSTESSSSSDSTSSQSSSTPTNNNSTTTPNNNTQOOSNTTPOQONN 660
QY 661 POPAOP 666
DB 661 POPAOP 666

RESULT 3
US-09-765-272-2

Sequence 2, Application US/09765272
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESSES:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue

CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-765-272-2

Query Match 100.0%; Score 666; DB 21; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KYDKNKNLIADLGSERRNAOANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLELN 60
DB 1 KYDKNKNLIADLGSERRNAOANDIPTDLVKAIVSIEDHREFDHGIDTIRILGAFLELN 60
QY 61 LOSNSLGGSTLTQOLIKLTFSTSTSDQTSIRKAOEAMLAIOLEKATKOEILTYIYNK 120
DB 61 LOSNSLGGSTLTQOLIKLTFSTSTSDQTSIRKAOEAMLAIOLEKATKOEILTYIYNK 120
QY 121 VYMSNGNTGMOTAONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPDEAADRRLVL 180
DB 121 VYMSNGNTGMOTAONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPDEAADRRLVL 180
QY 181 SEMKNOGYISAQOYERKAVNPTITDGLQSLKASNPAYMDNLYKEVINOVEETGYNLT 240
DB 181 SEMKNOGYISAQOYERKAVNPTITDGLQSLKASNPAYMDNLYKEVINOVEETGYNLT 240
QY 241 TGMDEVYTNDOAOKHLMIDYNTDEVVAYPPDELOVASTIVDSNGKVAIOLGARHOSN 300
DB 241 TGMDEVYTNDOAOKHLMIDYNTDEVVAYPPDELOVASTIVDSNGKVAIOLGARHOSN 300
QY 301 VSFGINQAVETNRDMSGTMRKPTTDYAPALEYGYVDSTATIVHDEPNYPGTNPVYNMNR 360
DB 301 VSFGINQAVETNRDMSGTMRKPTTDYAPALEYGYVDSTATIVHDEPNYPGTNPVYNMNR 360
QY 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFNLGIDYPSIHYSNAISSNTTESD 420
DB 361 GYFGNITLQYALQOOSRNPVAVETLNKVGINRAKTFNLGIDYPSIHYSNAISSNTTESD 420
QY 421 KKYGASSEKMAAAYAFANGGTYYKPMYTHKVFSDGSEKESNVTGRAMKETTAAYMMTD 480
DB 421 KKYGASSEKMAAAYAFANGGTYYKPMYTHKVFSDGSEKESNVTGRAMKETTAAYMMTD 480
QY 481 MKKTVLTGTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540
DB 481 MKKTVLTGTGRNAYLAWLPQAGKTGTSNYTDEIEIENHIKTSQFVAPDELFAGYTRKYSM 540

QY 541 AWTGYSNRLTPLYGNGLTVAAKVYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
| | | | |
DB 541 AWTGYSNRLTPLYGNGLTVAAKVYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
QY 601 STNNSPAPQOPPESTESSSDSSTSSSTSTSTNNSTTNPNNTQOSNTTPDOON 660
| | | | |
DB 601 STNNSPAPQOPPESTESSSDSSTSSSTSTSTNNSTTNPNNTQOSNTTPDOON 660
QY 661 POPAOP 666
| | | | |
DB 661 POPAOP 666
RESULT 4
US-09-583-110-3973
; Sequence 3973, Application US/09583110
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3973
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3973

Query Match 84.8%; Score 565; DB 19; Length 719;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 60
| | | | |
DB 54 KIYDNKNOLIADIGSERRNAQANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLRN 113
QY 61 LOSNSLQGGSTLTQOLIKLTYFSTSDOTISRRKQAEAMLAIOLEOKATKOELITYINK 120
| | | | |
DB 114 LOSNSLQGGSTLTQOLIKLTYFSTSDOTISRRKQAEAMLAIOLEOKATKOELITYINK 173
QY 121 VYWSNGNYGQTAQNYGKDLNLSLPOLATLALGMPQAFNOYDPYSHPEAODRRNLVL 180
| | | | |
DB 174 VYWSNGNYGQTAQNYGKDLNLSLPOLATLALGMPQAFNOYDPYSHPEAODRRNLVL 233
QY 181 SEKNNGYISAEQYKAVNTPITDGLQSLKASANTPAYMDNYLKEYINOVEEETGINLLT 240
| | | | |
DB 234 SEKNNGYISAEQYKAVNTPITDGLQSLKASANTPAYMDNYLKEYINOVEEETGINLLT 293
QY 241 TGMADVNTNDQEAOKHLMIDYNTDEVVAAPDDELQYASTIVDVSNGKVINOLGARHOSN 300
| | | | |
DB 294 TGMADVNTNDQEAOKHLMIDYNTDEVVAAPDDELQYASTIVDVSNGKVINOLGARHOSN 353
QY 301 VSFGINQAVETNRDMSGTMKPTIDYAPALEYGYVSTATTIVHDEPYNPPTNTPVYNMWR 360
| | | | |
DB 354 VSFGINQAVETNRDMSGTMKPTIDYAPALEYGYVSTATTIVHDEPYNPPTNTPVYNMWR 413
QY 361 GFPGNTTLOALQOSRNPVAVETLNVGLNRAKTFPLNGLGIDYPSIHYSAISSNTTESD 420
| | | | |
DB 414 GFPGNTTLOALQOSRNPVAVETLNVGLNRAKTFPLNGLGIDYPSIHYSAISSNTTESD 473
QY 421 KKYGASSEKMAAAYAFANGGTYYKPMYIHKYVSDGSEKEEENSGTRANKETATYMTD 480
| | | | |
DB 474 KKYGASSEKMAAAYAFANGGTYYKPMYIHKYVSDGSEKEEENSGTRANKETATYMTD 533

QY 481 MMKTVLYGTGRNAYLAMLPQAGKTGTSNTYDEIEHNIKTSQFVAPDELFAGYTRKYSM 540
| | | | |
DB 534 MMKTVLYGTGRNAYLAMLPQAGKTGTSNTYDEIEHNIKTSQFVAPDELFAGYTRKYSM 593
QY 541 AWTGYSNRLTPLYGNGLTVAAKVYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 600
| | | | |
DB 594 AWTGYSNRLTPLYGNGLTVAAKVYRSMNTYLSGSPNEDMNIPEGLYRNGEYFKNGAR 653
QY 601 STNNSPAPQOPPESTESSSDSSTSSSTSTSTNNSTTNPNNTQOSNTTPDOON 660
| | | | |
DB 654 STNNSPAPQOPPESTESSSDSSTSSSTSTSTNNSTTNPNNTQOSNTTPDOON 713
QY 661 POPAOP 666
| | | | |
DB 714 POPAOP 719

RESULT 5
US-09-107-433-3705
; Sequence 3705, Application US/09107433
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIA
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3705:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 721 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (b) LOCATION 1...721
; SEQUENCE DESCRIPTION: SEQ ID NO: 3705:
US-09-107-433-3705

Query Match 84.8%; Score 565; DB 15; Length 721;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 665; Conservative 0; Mismatches 1; Indels 0; Gaps 0;


```
QY 1 KIYDNKNOIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRPDRHGIDTIRILGAFLRN 60
DB 56 KIYDNKNOIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRPDRHGIDTIRILGAFLRN 115
QY 61 LOSNSILOGSSTLTQOLIKLTYFSTSDOTISKRAOEAMLAIOLEOKATKOEILTYINK 120
DB 116 LOSNSILOGSSTLTQOLIKLTYFSTSDOTISKRAOEAMLAIOLEOKATKOEILTYINK 175
QY 121 VYMSNGNYMOTPAONRYCKDLNNLSLPOLALLAGMPOANQYDPVSHPPAADRRNLVL 180
DB 176 VYMSNGNYMOTPAONRYCKDLNNLSLPOLALLAGMPOANQYDPVSHPPAADRRNLVL 235
QY 181 SEKNNGYISAEQYKAVNTPITDGLQSLKASNPAYMDNYLKEVYNQVEEETGNLTL 240
DB 236 SEKNNGYISAEQYKAVNTPITDGLQSLKASNPAYMDNYLKEVYNQVEEETGNLTL 295
QY 241 TGMDEVYTNVDOEAKHLMDIYNTDEVYAVPDELOVASTIVDSNGKVIAOLGARHOSN 300
DB 296 TGMDEVYTNVDOEAKHLMDIYNTDEVYAVPDELOVASTIVDSNGKVIAOLGARHOSN 355
QY 301 VSEGINQAVETNRDMSYMKPTDYAPALEGYVDSTATIVHDEPNYNGTNPVYNNMDR 360
DB 356 VSEGINQAVETNRDMSYMKPTDYAPALEGYVDSTATIVHDEPNYNGTNPVYNNMDR 415
QY 361 GFEGNTTLOVALQOOSRNVPAVETLNKVLNRAKTFNLGLGIDYPSIHYNSAISSNTTESD 420
DB 416 GFEGNTTLOVALQOOSRNVPAVETLNKVLNRAKTFNLGLGIDYPSIHYNSAISSNTTESD 475
QY 421 KYGASSEKMAAAYAFANGGTYKPMYIHKVYFSDGSEKESNVGTRAMKETTAYMMTD 480
DB 476 KYGASSEKMAAAYAFANGGTYKPMYIHKVYFSDGSEKESNVGTRAMKETTAYMMTD 535
QY 481 MKKTYVTYGRNAYLAMPQAGKTGTSNVTDEIEIENHKTQSOFVAPDELFAGYTRKYSN 540
DB 536 MKKTYVTYGRNAYLAMPQAGKTGTSNVTDEIEIENHKTQSOFVAPDELFAGYTRKYSN 595
QY 541 AVMTGYSNRLTPLYNGLTVAAKVYRSMNTYLSGSPNEDWNIPEGLYRNGEFVFKNGAR 600
DB 596 AVMTGYSNRLTPLYNGLTVAAKVYRSMNTYLSGSPNEDWNIPEGLYRNGEFVFKNGAR 655
QY 601 STWNSPAPQOPSTESSSSSSSDSSSTPSTNNSTTTNNNTTQOOSNTTPOOQONN 660
DB 656 STWNSPAPQOPSTESSSSSSSDSSSTPSTNNSTTTNNNTTQOOSNTTPOOQONN 715
QY 661 POPAOP 666
DB 716 POPAOP 721

RESULT 6
PCT-US02-03987-13423
; Sequence 13423, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elitra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits
; FILE REFERENCE: ELITRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US02-03987-13423
```

Query Match 49.5%; Score 330; DB 1; Length 719;
Best Local Similarity 99.6%; Pred. No. 0;

```
Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 72 LFOOLIKLTYFSTSDOTSRRKAOEAMLAIOLEOKATKOEILTYINKYMSNGNGMQ 131
DB 125 LFOOLIKLTYFSTSDOTSRRKAOEAMLAIOLEOKATKOEILTYINKYMSNGNGMQ 184
QY 132 TAAQNTYKXDLNNLSLPOLALLAGMPOANQYDPVSHPPAADRRNLVLEKNGYISA 191
DB 185 TAAQNTYKXDLNNLSLPOLALLAGMPOANQYDPVSHPPAADRRNLVLEKNGYISA 244
QY 192 EOYEKAVNTPITDGLQSLKASNPAYMDNYLKEVINQVEEETGNLTLTGMDEVYTNVDO 251
DB 245 EOYEKAVNTPITDGLQSLKASNPAYMDNYLKEVINQVEEETGNLTLTGMDEVYTNVDO 304
QY 252 EAKKHLMDIYNTDEVYAVPDELOVASTIVDSNGKVIAOLGARHOSNVSFGINQAVET 311
DB 305 EAKKHLMDIYNTDEVYAVPDELOVASTIVDSNGKVIAOLGARHOSNVSFGINQAVET 364
QY 312 NRDMSGTMRPTDYAPALEGYVDSTATIVHDEPNYNGTNPVYNNMDRGYFNGITLOYA 371
DB 365 NRDMSGTMRPTDYAPALEGYVDSTATIVHDEPNYNGTNPVYNNMDRGYFNGITLOYA 424
QY 372 LOOSRNVPAVETLNKVLNRAKTFNLGLGIDYPSIHYNSAISSNTTESDCKRYGASSEKMA 431
DB 425 LOOSRNVPAVETLNKVLNRAKTFNLGLGIDYPSIHYNSAISSNTTESDCKRYGASSEKMA 484
QY 432 AAYAAAFANGGTYKPMYIHKVYFSDGSEKESNVGTRAMKETTAYMMTDMMKTYVTLTGTG 491
DB 485 AAYAAAFANGGTYKPMYIHKVYFSDGSEKESNVGTRAMKETTAYMMTDMMKTYVTLTGTG 544
QY 492 RNAVYLA MLPQAGKTGTSNVTDEIEIENHKTQSOFVAPDELFAGYTRKYSMAVMTGYSNRLT 551
DB 545 RNAVYLA MLPQAGKTGTSNVTDEIEIENHKTQSOFVAPDELFAGYTRKYSMAVMTGYSNRLT 604
QY 552 PLYGNGLTVAAKVYRSMNTYLSGSPNEDWNIPEGLYRNGEFVFKNGARSTW 603
DB 605 PLYGNGLTVAAKVYRSMNTYLSGSPNEDWNIPEGLYRNGEFVFKNGARSTW 656

RESULT 7
US-09-815-242-13423
; Sequence 13423, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: Prokaryotes
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13423
```

LENGTH: 719
TYPE: PR
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13423

Query Match 49.5%; Score 330; DB 22; Length 719;

Best Local Similarity 99.6%; Pred. No. 0;
Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 72 LTQOLIKLTFSTSTQSTSRKAQEAFLAIQLEOKATKOEILTYINKYMSNGYGMQ 131
DB 125 LTQOLIKLTFSTSTQSTSRKAQEAFLAIQLEOKATKOEILTYINKYMSNGYGMQ 184
QY 132 TAAQNTYGGKDLNLSLPQALLAGMPAPNOYDPSHPPEAODRRNLVLSMKNOGYISA 191
DB 185 TAAQNTYGGKDLNLSLPQALLAGMPAPNOYDPSHPPEAODRRNLVLSMKNOGYISA 244
QY 192 EGYEKAVNPTITDGLQSLKASNPAYMDNYLKEVINQVEETGYNLTGMDVYTVNDQ 251
DB 245 EGYEKAVNPTITDGLQSLKASNPAYMDNYLKEVINQVEETGYNLTGMDVYTVNDQ 304
QY 252 EAQKHLMDIYNTDEYVAPDELOVASTIVDVSNGKYIAQIGARHOSNVSFGINQAVET 311
DB 305 EAQKHLMDIYNTDEYVAPDELOVASTIVDVSNGKYIAQIGARHOSNVSFGINQAVET 364
QY 312 NRDMGSTMKPITDYPAPALEGYVDSTATIVHDEPYNPGNTPYNMDRGYFGITLOYA 371
DB 365 NRDMGSTMKPITDYPAPALEGYVDSTATIVHDEPYNPGNTPYNMDRGYFGITLOYA 424
QY 372 LOOSRNPAYETLNKVLNRAKTFPLNGLGIDYPSIHSNAISSNTTESDKKYGASSEKMA 431
DB 425 LOOSRNPAYETLNKVLNRAKTFPLNGLGIDYPSIHSNAISSNTTESDKKYGASSEKMA 484
QY 432 AAYAAFPANGSTYKPKMYIHKVFSDSGESEKFSNNGTRAMETAYMMTDMMKTYLTYGTG 491
DB 485 AAYAAFPANGSTYKPKMYIHKVFSDSGESEKFSNNGTRAMETAYMMTDMMKTYLTYGTG 544
QY 492 RNAYLAWLPQAGKTGTSNTYDEIEENHIKTSQFAPAPDELFGAGYTRKXSMVWGYCSRMT 551
DB 545 RNAYLAWLPQAGKTGTSNTYDEIEENHIKTSQFAPAPDELFGAGYTRKXSMVWGYCSRMT 604
QY 552 PLYVNGLTVAAKYRSMATYLSGSPNEDMNIPEGLYRNGEFYFKNGARSTW 603
DB 605 PLYVNGLTVAAKYRSMATYLSGSPNEDMNIPEGLYRNGEFYFKNGARSTW 656

RESULT 8

US-10-072-851-13423
Sequence 13423, Application US/10072851

GENERAL INFORMATION:

APPLICANT: Carr, Grant J.
APPLICANT: Xu, H. Howard
APPLICANT: Foulkes, J. Gordon
APPLICANT: Zamudio, Carlos
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Roemer, Terry
APPLICANT: Jiang, Bo
APPLICANT: Boone, Charles
APPLICANT: Bussey, Howard
TITLE OF INVENTION: Methods for identifying the target of a compound which inhibits
FILE REFERENCE: ELITRA.028A
CURRENT APPLICATION NUMBER: US/10/072,851
CURRENT FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
NUMBER OF SEQ ID NOS: 15811

SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13423
LENGTH: 719
TYPE: PR
ORGANISM: Streptococcus pneumoniae
US-10-072-851-13423

Query Match 49.5%; Score 330; DB 24; Length 719;

Best Local Similarity 99.6%; Pred. No. 0;
Matches 530; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 72 LTQOLIKLTFSTSTQSTSRKAQEAFLAIQLEOKATKOEILTYINKYMSNGYGMQ 131
DB 125 LTQOLIKLTFSTSTQSTSRKAQEAFLAIQLEOKATKOEILTYINKYMSNGYGMQ 184
QY 132 TAAQNTYGGKDLNLSLPQALLAGMPAPNOYDPSHPPEAODRRNLVLSMKNOGYISA 191
DB 185 TAAQNTYGGKDLNLSLPQALLAGMPAPNOYDPSHPPEAODRRNLVLSMKNOGYISA 244
QY 192 EGYEKAVNPTITDGLQSLKASNPAYMDNYLKEVINQVEETGYNLTGMDVYTVNDQ 251
DB 245 EGYEKAVNPTITDGLQSLKASNPAYMDNYLKEVINQVEETGYNLTGMDVYTVNDQ 304
QY 252 EAQKHLMDIYNTDEYVAPDELOVASTIVDVSNGKYIAQIGARHOSNVSFGINQAVET 311
DB 305 EAQKHLMDIYNTDEYVAPDELOVASTIVDVSNGKYIAQIGARHOSNVSFGINQAVET 364
QY 312 NRDMGSTMKPITDYPAPALEGYVDSTATIVHDEPYNPGNTPYNMDRGYFGITLOYA 371
DB 365 NRDMGSTMKPITDYPAPALEGYVDSTATIVHDEPYNPGNTPYNMDRGYFGITLOYA 424
QY 372 LOOSRNPAYETLNKVLNRAKTFPLNGLGIDYPSIHSNAISSNTTESDKKYGASSEKMA 431
DB 425 LOOSRNPAYETLNKVLNRAKTFPLNGLGIDYPSIHSNAISSNTTESDKKYGASSEKMA 484
QY 432 AAYAAFPANGSTYKPKMYIHKVFSDSGESEKFSNNGTRAMETAYMMTDMMKTYLTYGTG 491
DB 485 AAYAAFPANGSTYKPKMYIHKVFSDSGESEKFSNNGTRAMETAYMMTDMMKTYLTYGTG 544
QY 492 RNAYLAWLPQAGKTGTSNTYDEIEENHIKTSQFAPAPDELFGAGYTRKXSMVWGYCSRMT 551
DB 545 RNAYLAWLPQAGKTGTSNTYDEIEENHIKTSQFAPAPDELFGAGYTRKXSMVWGYCSRMT 604
QY 552 PLYVNGLTVAAKYRSMATYLSGSPNEDMNIPEGLYRNGEFYFKNGARSTW 603
DB 605 PLYVNGLTVAAKYRSMATYLSGSPNEDMNIPEGLYRNGEFYFKNGARSTW 656

RESULT 9

US-08-116-541-4
Sequence 4, Application US/08116541

GENERAL INFORMATION:

APPLICANT: Masure, H. Robert
APPLICANT: Pearce, Barbara J.
APPLICANT: Toumanen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/116,541

FILING DATE: 19930901
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-069
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-116-541-4

Query Match 36.8%; Score 245; DB 5; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.6e-232; Indels 0; Gaps 0;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 LTQQLIKLYFSTSDQTSKRAQEWLAIQLEKATKQELITYYINKVYNSNGNYGMQ 131
DB 76 LTQQLIKLYFSTSDQTSKRAQEWLAIQLEKATKQELITYYINKVYNSNGNYGMQ 135
QY 132 TAAQNTYGGDNLNLSPQLALLAGMPQAPNOVDPSHPRAQDRRLVLEKMNQYISA 191
DB 136 TAAQNTYGGDNLNLSPQLALLAGMPQAPNOVDPSHPRAQDRRLVLEKMNQYISA 195
QY 192 EGYEKAVNPITDGLQSLKSNYPAYMDNYLKEVINQVEETGYMLTTGMDVYTNVDO 251
DB 196 EGYEKAVNPITDGLQSLKSNYPAYMDNYLKEVINQVEETGYMLTTGMDVYTNVDO 255
QY 252 EAQKHLMDIYNTDEVYAYPDDELQVASTIVDVSNGKVIQAQLGARHSSNVSEGINQAVET 311
DB 256 EAQKHLMDIYNTDEVYAYPDDELQVASTIVDVSNGKVIQAQLGARHSSNVSEGINQAVET 315
QY 312 NRDMG 316
DB 316 NRDMG 320

RESULT 10

US-08-116-541-24
Sequence 24, Application US/08116541
GENERAL INFORMATION:
APPLICANT: Masure, H. Robert
APPLICANT: Pearce, Barbara J.
APPLICANT: Toumanen, Elaine
TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
TITLE OF INVENTION: ACCELLULAR VACCINES BASED THERON
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/116,541
FILING DATE: 19930901
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-069
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
STRAIN: R6
IMMEDIATE SOURCE:
CLONE: SPRU42
US-08-116-541-24

Query Match 11.6%; Score 77; DB 5; Length 77;
Best Local Similarity 100.0%; Pred. No. 3.6e-67;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 TTGMDVYTNVDOEAQKHLMDIYNTDEVYAYPDDELQVASTIVDVSNGKVIQAQLGARHSS 299
DB 1 TTGMDVYTNVDOEAQKHLMDIYNTDEVYAYPDDELQVASTIVDVSNGKVIQAQLGARHSS 60
QY 300 NVSEGINQAVETNRDMG 316
DB 61 NVSEGINQAVETNRDMG 77

RESULT 11

US-09-107-532-5667
Sequence 5667, Application US/09107532
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7308
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER:
OPERATING SYSTEM:
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532
FILING DATE:
APPLICATION DATA:
APPLICATION NUMBER: 60/085598
FILING DATE: May 14, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5667:

SEQUENCE CHARACTERISTICS:
LENGTH: 823 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1...823
US-09-107-532-5667

Query Match 2.7%; Score 18; DB 15; Length 823;
Best Local Similarity 100.0%; Pred. No. 6 6e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 429 KMAAAYAFANGTYRKP 446
|||||
DB 543 KMAAAYAFANGTYRKP 560

RESULT 12
US-09-107-532A-5667
; Sequence 5667, Application US/09107532A
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-8277
; TELEFAX: (781)893-5507
; INFORMATION FOR SEQ ID NO: 5667:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 823 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (8) LOCATION 1...823
; SEQUENCE DESCRIPTION: SEQ ID NO: 5667:
US-09-107-532A-5667

Query Match 2.7%; Score 18; DB 15; Length 823;
Best Local Similarity 100.0%; Pred. No. 6 6e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 429 KMAAAYAFANGTYRKP 446
|||||
DB 543 KMAAAYAFANGTYRKP 560

RESULT 13
US-09-634-238-229
; Sequence 229, Application US/09634238
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.104301
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 229
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-229

Query Match 2.1%; Score 14; DB 20; Length 771;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 LOGGSTLTQOLIKL 79
|||||
DB 143 LOGGSTLTQOLIKL 156

RESULT 14
US-09-634-238-262
; Sequence 262, Application US/09634238
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bloksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna C.
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides, materials incorporating
; TITLE OF INVENTION: them and methods for using them.
; FILE REFERENCE: 11000.104301
; CURRENT APPLICATION NUMBER: US/09/634,238
; CURRENT FILING DATE: 2000-08-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 262
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-262

```
Query Match          2.1%; Score 14; DB 20; Length 771;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LOGGSTLTQQLIKL 79
   |||||
DB 143 LOGGSTLTQQLIKL 156

RESULT 15
PCT-US02-03987-10728
; Sequence 10728, Application PC/TUS0203987
; GENERAL INFORMATION:
; APPLICANT: Elittra Pharmaceuticals, Inc.
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibits C
; FILE REFERENCE: ELITTRA.028VPC
; CURRENT APPLICATION NUMBER: PCT/US02/03987
; CURRENT FILING DATE: 2002-02-02
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
PCT-US02-03987-10728

Query Match          2.1%; Score 14; DB 1; Length 778;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LOGGSTLTQQLIKL 79
   |||||
DB 146 LOGGSTLTQQLIKL 159

RESULT 16
US-09-815-242-10728
; Sequence 10728, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITTRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10728

Query Match          2.1%; Score 14; DB 22; Length 778;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LOGGSTLTQQLIKL 79
   |||||
DB 146 LOGGSTLTQQLIKL 159

RESULT 17
US-10-072-851-10728
; Sequence 10728, Application US/10072851
; GENERAL INFORMATION:
; APPLICANT: Carr, Grant J.
; APPLICANT: Xu, H. Howard
; APPLICANT: Foulkes, J. Gordon
; APPLICANT: Zamudio, Carlos
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Methods for Identifying the Target of a Compound which Inhibit
; FILE REFERENCE: ELITTRA.028A
; CURRENT APPLICATION NUMBER: US/10/072,851
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 15811
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10728
; LENGTH: 778
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-10-072-851-10728

Query Match          2.1%; Score 14; DB 24; Length 778;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 LOGGSTLTQQLIKL 79
   |||||
DB 146 LOGGSTLTQQLIKL 159

RESULT 18
US-09-134-000-4939
; Sequence 4939, Application US/09134000A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
; FILE REFERENCE: GTC-005
; CURRENT APPLICATION NUMBER: US/09/134,000A
; CURRENT FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 6810
; SEQ ID NO 4939
; LENGTH: 789
```

TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000-4939

Query Match 2.1%; Score 14; DB 15; Length 789;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 GGSTLTQOLIK 79
Db 157 GGSTLTQOLIK 170

RESULT 19
PCT-US97-14436-535
Sequence 535, Application PC/TUS9714436
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/14436
FILING DATE: 15-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 535:
SEQUENCE CHARACTERISTICS:
LENGTH: 266 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US97-14436-535

Query Match 1.7%; Score 11; DB 1; Length 266;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
Db 107 GGSTLTQOLIK 117

RESULT 20
US-08-911-503-535
Sequence 535, Application US/08911503
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,503
FILING DATE: 15-AUG-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 535:
SEQUENCE CHARACTERISTICS:
LENGTH: 266 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-911-503-535

Query Match 1.7%; Score 11; DB 13; Length 266;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSTLTQOLIK 78
Db 107 GGSTLTQOLIK 117

RESULT 21
US-08-911-503A-535
Sequence 535, Application US/08911503A
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
NUMBER OF SEQUENCES: 713
CORRESPONDENCE ADDRESS:

ADDRESS: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911.503A
FILING DATE: 15-AUG-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,022
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50533
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 535:
SEQUENCE CHARACTERISTICS:
LENGTH: 266 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-911-503A-535

Query Match 1.7%; Score 11; DB 13; Length 266;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSITLQOLIK 78
|||||
DB 107 GGSITLQOLIK 117

RESULT 22
US-09-417-507-35782
Sequence 35782, Application US/09417507
GENERAL INFORMATION:
APPLICANT: KEITH G. WEINSTOCK ET AL.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
FILE REFERENCE: PAT99-10
CURRENT APPLICATION NUMBER: US/09/417.507
FILING DATE: 1999-10-14
NUMBER OF SEQ ID NOS: 44312
SEQ ID NO 35782
LENGTH: 314
TYPE: PRT
ORGANISM: A.fumigatus
US-09-417-507-35782

Query Match 1.7%; Score 11; DB 18; Length 314;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSSTS 626
|||||
DB 148 SSSSSSDSSTS 158

RESULT 23

US-09-583-110-4219
Sequence 4219, Application US/09583110
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococci
FILE REFERENCE: PAT00-07A
CURRENT APPLICATION NUMBER: US/09/583.110
FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 4219
LENGTH: 820
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-4219

Query Match 1.7%; Score 11; DB 19; Length 820;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSITLQOLIK 78
|||||
DB 176 GGSITLQOLIK 186

RESULT 24
US-09-377-465-2
Sequence 2, Application US/09377465
GENERAL INFORMATION:
APPLICANT: Hoskins, JoAnn
APPLICANT: Jaskunas, S. Richard
APPLICANT: Zhao, Genshi
APPLICANT: Rocky, Pamela
TITLE OF INVENTION: Novel Penicillin Binding Protein From Streptococcus
FILE REFERENCE: X12498
CURRENT APPLICATION NUMBER: US/09/377.465
FILING DATE: 1999-08-19
EARLIER APPLICATION NUMBER: 60/100,887
EARLIER FILING DATE: 1998-09-23
EARLIER APPLICATION NUMBER: 60/111,862
EARLIER FILING DATE: 1998-12-11
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 821
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-377-465-2

Query Match 1.7%; Score 11; DB 17; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSITLQOLIK 78
|||||
DB 176 GGSITLQOLIK 186

RESULT 25
US-09-377-465A-2
Sequence 2, Application US/09377465A
GENERAL INFORMATION:
APPLICANT: Hoskins, JoAnn
APPLICANT: Jaskunas, Stanley R

APPLICANT: Zhao, Genshi
APPLICANT: Rockey, Pamela K
TITLE OF INVENTION: NOVEL PENICILLIN BINDING PROTEIN FROM STREPTOCOCCUS
TITLE OF INVENTION: PNEUMONIAE
FILE REFERENCE: X12498 Sequence List
CURRENT APPLICATION NUMBER: US/09/377,465A
CURRENT FILING DATE: 1999-08-19
PRIOR APPLICATION NUMBER: 60/100,887
PRIOR FILING DATE: 1998-09-23
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 821
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-377-465A-2

Query Match 1.7%; Score 11; DB 17; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSSTLQOLIK 78
|||||
DB 176 GGSSTLQOLIK 186

RESULT 26
US-09-107-433-4760

Sequence 4760, Application US/09107433
GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS FOR DIAGN

NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Atinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4760:
SEQUENCE CHARACTERISTICS:
LENGTH: 834 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:

NAME/KEY: misc-feature
LOCATION: (B) LOCATION 1...834
SEQUENCE DESCRIPTION: SEQ ID NO: 4760:
US-09-107-433-4760

Query Match 1.7%; Score 11; DB 15; Length 834;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GGSSTLQOLIK 78
|||||
DB 190 GGSSTLQOLIK 200

RESULT 27
US-09-621-976-7706

Sequence 7706, Application US/09621976
GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 7706

LENGTH: 76

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: UNSURE

LOCATION: 72

OTHER INFORMATION: Xaa = Ala, Pro, Ser, Thr
US-09-621-976-7706

Query Match 1.5%; Score 10; DB 20; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
|||||
DB 23 SSSSSSDSST 32

RESULT 28
US-09-834-366-24929

Sequence 24929, Application US/09834366

GENERAL INFORMATION:

APPLICANT: Bejani, Stephane

APPLICANT: Tanaka, Hiroaki

APPLICANT: Dumas Milne Edwards, Jean Baptiste

APPLICANT: Jobert, Severin

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: 81.US2.REG

CURRENT APPLICATION NUMBER: US/09/834,366

CURRENT FILING DATE: 2001-04-13

PRIOR APPLICATION NUMBER: US 60/197,873

PRIOR FILING DATE: 2000-04-18

NUMBER OF SEQ ID NOS: 52153

SOFTWARE: Patent.pm

SEQ ID NO 24929

LENGTH: 76

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: UNSURE

LOCATION: 72

OTHER INFORMATION: Xaa = Ala, Pro, Ser, Thr

US-09-834-366-24929

Query Match 1.5%; Score 10; DB 22; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625

DB 23 SSSSSSDSST 32

RESULT 29

US-09-834-366-24930

; Sequence 24930, Application US/09834366

; GENERAL INFORMATION:

; APPLICANT: Bejanin, Stephane

; APPLICANT: Tanaka, Hiroaki

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Jobert, Severin

; APPLICANT: Giordano, Jean-Yves

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: 81.052.REG

; CURRENT APPLICATION NUMBER: US/09/834,366

; CURRENT FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: US 60/197,873

; PRIOR FILING DATE: 2000-04-18

; NUMBER OF SEQ ID NOS: 52153

; SOFTWARE: Patent.pm

; SEQ ID NO 24930

; LENGTH: 76

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: UNSURE

; LOCATION: 72

; OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr

US-09-834-366-24930

Query Match 1.5%; Score 10; DB 22; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625

DB 23 SSSSSSDSST 32

RESULT 30

US-60-147-499-7706

; Sequence 7706, Application US/60147499

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Jobert, S.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/60/147,499

; CURRENT FILING DATE: 1999-08-05

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 7706

; LENGTH: 76

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: UNSURE

; LOCATION: 72

; OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr

US-60-147-499-7706

Query Match 1.5%; Score 10; DB 26; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625

DB 23 SSSSSSDSST 32

RESULT 31

US-60-197-873-24929

; Sequence 24929, Application US/60197873

; GENERAL INFORMATION:

; APPLICANT: Bejanin, Stephane

; APPLICANT: Tanaka, Hiroaki

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: 81.US1.PRO

; CURRENT APPLICATION NUMBER: US/60/197,873

; CURRENT FILING DATE: 2000-04-18

; NUMBER OF SEQ ID NOS: 52153

; SOFTWARE: Patent.pm

; SEQ ID NO 24929

; LENGTH: 76

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: UNSURE

; LOCATION: 72

; OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr

US-60-197-873-24929

Query Match 1.5%; Score 10; DB 26; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625

DB 23 SSSSSSDSST 32

RESULT 32

US-60-197-873-24930

; Sequence 24930, Application US/60197873

; GENERAL INFORMATION:

; APPLICANT: Bejanin, Stephane

; APPLICANT: Tanaka, Hiroaki

; APPLICANT: Dumas Milne Edwards, Jean Baptiste

; APPLICANT: Jobert, Severin

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: 81.US1.PRO

; CURRENT APPLICATION NUMBER: US/60/197,873

; CURRENT FILING DATE: 2000-04-18

; NUMBER OF SEQ ID NOS: 52153

; SOFTWARE: Patent.pm

; SEQ ID NO 24930

; LENGTH: 76

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: UNSURE

; LOCATION: 72

; OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr

US-60-197-873-24930

Query Match 1.5%; Score 10; DB 26; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.55;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
Db 23 SSSSSSDSST 32

RESULT 33

US-09-536-784-102
; Sequence 102, Application US/09536784
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-09-536-784-102

Query Match 1.5%; Score 10; DB 19; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
Db 18 SSSSSSDSST 27

RESULT 34

US-09-765-271-102
; Sequence 102, Application US/09765271
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,271
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/536,784
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-09-765-271-102

Query Match 1.5%; Score 10; DB 21; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
Db 18 SSSSSSDSST 27

RESULT 35

US-09-765-272-102
; Sequence 102, Application US/09765272
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504

```
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-09-765-272-102
```

```
Query Match      1.5%; Score 10; DB 21; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 616 SSSSSSDSST 625
      |||
Db 18 SSSSSSDSST 27
```

```
RESULT 36
US-09-614-150-34311
: Sequence 34311, Application US/09614150
: GENERAL INFORMATION:
: APPLICANT: Venter, J. Craig
: APPLICANT: et al.
: TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
: TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
: FILE REFERENCE: CLO00728
: CURRENT APPLICATION NUMBER: US/09/614,150
: CURRENT FILING DATE: 2000-07-11
: PRIOR APPLICATION NUMBER: 60/157,832
: PRIOR FILING DATE: 1999-10-05
: PRIOR APPLICATION NUMBER: 60/160,191
: PRIOR FILING DATE: 1999-10-19
: PRIOR APPLICATION NUMBER: 60/161,932
: PRIOR FILING DATE: 1999-10-28
: PRIOR APPLICATION NUMBER: 60/164,769
: PRIOR FILING DATE: 1999-11-12
: PRIOR APPLICATION NUMBER: 60/173,383
: PRIOR FILING DATE: 1999-12-28
: PRIOR APPLICATION NUMBER: 60/175,693
: PRIOR FILING DATE: 2000-01-12
: PRIOR APPLICATION NUMBER: 60/184,831
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: 60/191,637
: PRIOR FILING DATE: 2000-03-23
: NUMBER OF SEQ ID NOS: 43008
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 34311
: LENGTH: 119
: TYPE: PRT
: ORGANISM: DROSOPHILA
US-09-614-150-34311
```

```
Query Match      1.5%; Score 10; DB 20; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 617 SSSSSSDSST 626
      |||
Db 60 SSSSSSDSST 69
```

```
RESULT 37
US-60-191-637-33895
: Sequence 33895, Application US/60191637
: GENERAL INFORMATION:
: APPLICANT: Venter, J. Craig
: TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
```

```
: TITLE OF INVENTION: GENE SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: CLO00392
: CURRENT APPLICATION NUMBER: US/60/191,637
: CURRENT FILING DATE: 2000-03-23
: NUMBER OF SEQ ID NOS: 42660
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 33895
: LENGTH: 119
: TYPE: PRT
: ORGANISM: DROSOPHILA
US-60-191-637-33895
```

```
Query Match      1.5%; Score 10; DB 26; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.84;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 617 SSSSSSDSST 626
      |||
Db 60 SSSSSSDSST 69
```

```
RESULT 38
US-09-621-976-5866
: Sequence 5866, Application US/09621976
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J.B.
: APPLICANT: Jobert, S.
: APPLICANT: Giordano, J.Y.
: TITLE OF INVENTION: ESTS and Encoded Human Proteins.
: FILE REFERENCE: GENSET.054PR2
: CURRENT APPLICATION NUMBER: US/09/621,976
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 19335
: SOFTWARE: Patent.pm
: SEQ ID NO 5866
: LENGTH: 130
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SIGNAL
: LOCATION: -36...-1
: NAME/KEY: UNSURE
: LOCATION: 90
: OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr
US-09-621-976-5866
```

```
Query Match      1.5%; Score 10; DB 20; Length 130;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 616 SSSSSSDSST 625
      |||
Db 77 SSSSSSDSST 86
```

```
RESULT 39
US-60-147-499-5866
: Sequence 5866, Application US/60147499
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J.B.
: APPLICANT: Jobert, S.
: APPLICANT: Giordano, J.Y.
: TITLE OF INVENTION: ESTS and Encoded Human Proteins.
: FILE REFERENCE: GENSET.054PR2
: CURRENT APPLICATION NUMBER: US/60/147,499
: CURRENT FILING DATE: 1999-08-05
: NUMBER OF SEQ ID NOS: 19335
: SOFTWARE: Patent.pm
: SEQ ID NO 5866
: LENGTH: 130
```

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: -36..-1
FEATURE:
NAME/KEY: UNSURE
LOCATION: 90
OTHER INFORMATION: Xaa = Ala,Pro,Ser,Thr
US-60-147-499-5866

Query Match 1.5%; Score 10; DB 26; Length 130;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 77 SSSSSSDSST 86

RESULT 40
US-09-950-083-3079
Sequence 3079, Application US/09950083
GENERAL INFORMATION:
APPLICANT: Rosen, et. al
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS805
CURRENT APPLICATION NUMBER: US/09/950,083
CURRENT FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: 60/278,650
PRIOR FILING DATE: 2001-03-27
PRIOR APPLICATION NUMBER: 09/833,245
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: PCT/US01/11988
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: PCT/US00/06043
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06012
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06058
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06044
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06059
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06042
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06014
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06013
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06049
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06057
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: PCT/US00/06824
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06765
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06792
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06830
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06782
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06822
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06791
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06828
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06823

PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/06781
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: PCT/US00/07505
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07440
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07506
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07507
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07535
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07525
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07534
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07483
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07526
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07527
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: PCT/US00/07661
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07579
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07723
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07724
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/14929
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/07722
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07578
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07726
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07677
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/07725
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/US00/09070
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08982
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08983
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09067
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09066
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09068
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08981
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/08980
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09071
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/09069
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/US00/15136
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14926
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14963
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/15135
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: PCT/US00/14934
PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: PCT/US00/14933
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/15137
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14928
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14973
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/14964
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/26376
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26371
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26324
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26323
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US00/26337
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: PCT/US01/13318
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/124,146
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/167,061
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/124,093
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/166,989
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/124,145
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,654
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,099
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,661
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,096
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,622
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,143
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/168,663
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,095
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/138,598
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,665
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/125,360
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: US 60/138,626
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,662
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,144
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/138,574
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,667
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/124,142
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: US 60/138,597
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: US 60/168,666
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/125,359
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: US 60/168,664

; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: US 60/126,051
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/169,906

Query Match 1.5%; Score 10; DB 23; Length 194;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 149 SSSSSSDSST 158

RESULT 41
PCT-US01-27760-504
; Sequence 504, Application PC/TUS0127760
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-115/785
; CURRENT APPLICATION NUMBER: PCT/US01/27760
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/687,527
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 892
; SOFTWARE: Custom
; SEQ ID NO 504
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-27760-504

Query Match 1.5%; Score 10; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 218 SSSSSSDSST 227

RESULT 42
PCT-US01-27760A-504
; Sequence 504, Application PC/TUS0127760A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-115/785
; CURRENT APPLICATION NUMBER: PCT/US01/27760A
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/687,527
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 892
; SOFTWARE: Custom
; SEQ ID NO 504
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-27760A-504

Query Match 1.5%; Score 10; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 218 SSSSSSDSST 227

RESULT 43

US-09-790-893-6
; Sequence 6, Application US/09790893
; GENERAL INFORMATION:
; APPLICANT: Mauck, Kimberly
; APPLICANT: Bush, Angie
; APPLICANT: Heichman, Karen
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: 2318-293
; CURRENT APPLICATION NUMBER: US/09/790,893
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: US 60/185,080
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-790-893-6

Query Match 1.5%; Score 10; DB 21; Length 263;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
Db 218 SSSSSSDSST 227

RESULT 44
US-09-056-019-3
; Sequence 3, Application US/09056019A
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wizemann, Theresa
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Koenig, Scott
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; FILE REFERENCE: 1340-1-017
; CURRENT APPLICATION NUMBER: US/09/056,019A
; CURRENT FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-056-019-3

Query Match 1.5%; Score 10; DB 14; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
Db 113 SSSSSSDSST 122

RESULT 45
US-09-287-788-3
; Sequence 3, Application US/09287788B
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Masure, H. R.
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL

; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; FILE REFERENCE: 1340-1-017N
; CURRENT APPLICATION NUMBER: US/09/287,788B
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: 60/080,878
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-287-788-3

Query Match 1.5%; Score 10; DB 16; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
Db 113 SSSSSSDSST 122

RESULT 46
US-09-056-019-1
; Sequence 1, Application US/09056019A
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Wizemann, Theresa
; APPLICANT: Masure, H. R.
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Koenig, Scott
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; FILE REFERENCE: 1340-1-017
; CURRENT APPLICATION NUMBER: US/09/056,019A
; CURRENT FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-056-019-1

Query Match 1.5%; Score 10; DB 14; Length 406;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 SSSSSSDSST 625
Db 113 SSSSSSDSST 122

RESULT 47
US-09-287-788-1
; Sequence 1, Application US/09287788B
; GENERAL INFORMATION:
; APPLICANT: Tuomanen, Elaine I
; APPLICANT: Masure, H. R.
; TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
; TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
; FILE REFERENCE: 1340-1-017N
; CURRENT APPLICATION NUMBER: US/09/287,788B
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: 60/080,878
; PRIOR FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 39

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 406
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-287-788-1

Query Match 1.5%; Score 10; DB 16; Length 406;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 113 SSSSSSDSST 122

RESULT 48
US-09-056-019-24
Sequence 24, Application US/09056019A
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine I
APPLICANT: Witzemann, Theresa
APPLICANT: Masure, H. R.
APPLICANT: Johnson, Leslie S.
TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
TITLE OF INVENTION: THEREFROM AND USES THEREOF
FILE REFERENCE: 1340-1-017
CURRENT APPLICATION NUMBER: US/09/056, 019A
CURRENT FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 428
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-056-019-24

Query Match 1.5%; Score 10; DB 14; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 113 SSSSSSDSST 122

RESULT 49
US-09-287-788-24
Sequence 24, Application US/09287788B
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine I
APPLICANT: Masure, H. R.
TITLE OF INVENTION: POLYPEPTIDE COMPRISING THE AMINO ACID OF AN N-TERMINAL
TITLE OF INVENTION: CHOLINE BINDING PROTEIN A TRUNCATE, VACCINE DERIVED
TITLE OF INVENTION: THEREFROM AND USES THEREOF
FILE REFERENCE: 1340-1-017N
CURRENT APPLICATION NUMBER: US/09/287, 788B
CURRENT FILING DATE: 1999-04-07
PRIOR APPLICATION NUMBER: 60/080, 878
PRIOR FILING DATE: 1998-04-07
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 428
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-287-788-24

Query Match 1.5%; Score 10; DB 16; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 113 SSSSSSDSST 122

RESULT 50
US-09-605-703B-1256
Sequence 1256, Application US/09605703B
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Krieger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
TITLE OF INVENTION: PROTEINS
FILE REFERENCE: BGI-129CP
CURRENT APPLICATION NUMBER: US/09/605, 703B
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: 60/142, 764
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 60/152, 318
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 2934
SEQ ID NO 1256
LENGTH: 428
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-605-703B-1256

Query Match 1.5%; Score 10; DB 20; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
|||||
DB 59 SSSSSSDSST 68

Search completed: June 13, 2002, 08:48:47
Job time: 243 sec

• • • • •

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:45:49 ; Search time 20.8 Seconds
(without alignments)
2561.523 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 666
Sequence: 1 KIYDNKNQIADIGSERRV.....TQSSNTTPDOQNONPQAP 666

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 244551 seqs, 7999522 residues

Word size : 10

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Pending_Patents_AA_New:*

1: /cgn2_6/ptodata/1/paa/PCr_NEW.COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW.COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description
	1	666	100.0	666	4	US-08-961-083-2	Sequence 2, Appli
	2	11	1.7	821	5	US-09-769-744A-122	Sequence 122, App
	3	10	1.5	110	4	US-08-961-083-102	Sequence 102, App
	4	10	1.5	194	6	US-10-105-299-4083	Sequence 4083, Ap
	5	10	1.5	602	7	US-60-360-039-3150	Sequence 3150, Ap
	6	10	1.5	1704	5	US-09-991-262-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-08-961-083-2
; Sequence 2, Application US/08961083

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA: US/08/961,083
APPLICATION NUMBER: US/08/961,083
FILING DATE: 30-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-961-083-2

Query Match 100.0%; Score 666; DB 4; Length 666;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	KIYDNKNQIADIGSERRVNAQNDIPTDLVKAIVSIEDHFFPDHNGIDTIRILGAFLRN	60
DB	1	KIYDNKNQIADIGSERRVNAQNDIPTDLVKAIVSIEDHFFPDHNGIDTIRILGAFLRN	60
QY	61	LOSNSILOGSSTLTQOLIKITFYESTSTSDOTISKRAQEAHLAIQLEQATKQELITYYINK	120
DB	61	LOSNSILOGSSTLTQOLIKITFYESTSTSDOTISKRAQEAHLAIQLEQATKQELITYYINK	120
QY	121	VYMSNGNYMGQTAQNYGYKDLNLSLPQLALLAGMPQAPNOYDPYSHPEAADRNRNLYL	180
DB	121	VYMSNGNYMGQTAQNYGYKDLNLSLPQLALLAGMPQAPNOYDPYSHPEAADRNRNLYL	180
QY	181	SEKKNNGYISAEQYKAVNTPITDGLQSLKASNTYAYMDNYLKEYINOVEETGYNLLT	240
DB	181	SEKKNNGYISAEQYKAVNTPITDGLQSLKASNTYAYMDNYLKEYINOVEETGYNLLT	240
QY	241	TGMDVYTNVDOEKHLMIDYNTDEYVAYPDELAQVASTIVDSNGKYAQLGARQSSN	300
DB	241	TGMDVYTNVDOEKHLMIDYNTDEYVAYPDELAQVASTIVDSNGKYAQLGARQSSN	300
QY	301	VSEGINQAVETNRDWSGTYMKPTTDYAPALEYGYVDSATIVHDEPNYPCGTNPVYNNMR	360
DB	301	VSEGINQAVETNRDWSGTYMKPTTDYAPALEYGYVDSATIVHDEPNYPCGTNPVYNNMR	360
QY	361	GYFGNTTIOYALQOSNRNVAVETLNKVGILNRATFTLNGLGIDVPSIHYSAISSNTTESD	420
DB	361	GYFGNTTIOYALQOSNRNVAVETLNKVGILNRATFTLNGLGIDVPSIHYSAISSNTTESD	420
QY	421	KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYFSGSSEKESFNVTGRAMKETTAAMMD	480
DB	421	KRYGASSEKMAAAYAFANGGTYYKPMYIHKVYFSGSSEKESFNVTGRAMKETTAAMMD	480
QY	481	MMKTYVLTGTGRNAYLAWLPQAGKTGTSNYTDEIEENHIKTSQFVAPDELFACTRRYSM	540
DB	481	MMKTYVLTGTGRNAYLAWLPQAGKTGTSNYTDEIEENHIKTSQFVAPDELFACTRRYSM	540
QY	541	AWMTGYSNRLTPLYGNGTLVAAKVVYSMMTYLSEGSNPDMDNIPBELYRNGEVRFNNGAR	600
DB	541	AWMTGYSNRLTPLYGNGTLVAAKVVYSMMTYLSEGSNPDMDNIPBELYRNGEVRFNNGAR	600

QY 601 STWNSAPDQPPSTESSSSSDSTSSSTPSTNNSTTNPNNNTQOSNTTPDOON 660
DB 601 STWNSAPDQPPSTESSSSSDSTSSSTPSTNNSTTNPNNNTQOSNTTPDOON 660
QY 661 POPAP 666
DB 661 POPAP 666

RESULT 2

US-09-769-744A-122
; Sequence 122, Application US/09769744A
; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PMC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 122
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-122

Query Match 1.7%; Score 11; DB 5; Length 821;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GGSTLTQOLIK 78
DB 176 GGSTLTQOLIK 186

RESULT 3

US-08-961-083-102
; Sequence 102, Application US/08961083
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-08-961-083-102

Query Match 1.5%; Score 10; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 18 SSSSSSDSST 27

RESULT 4

US-10-105-299-4083
; Sequence 4083, Application US/10105299
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS950
; CURRENT APPLICATION NUMBER: US/10/105,299
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 15197
; Prior Application removed. See File Wrapper or Palm
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO: 4083
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-299-4083

Query Match 1.5%; Score 10; DB 6; Length 194;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 SSSSSSDSST 625
DB 149 SSSSSSDSST 158

RESULT 5

US-60-360-039-3150
; Sequence 3150, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO: 3150
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:

NAME/KEY: unsure
LOCATION: (1)...(602)
OTHER INFORMATION: unsure at all Xaa locations
US-60-360-039-3150

Query Match 1.5%; Score 10; DB 7; Length 602;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 615 ESSSSSDSS 624
Db 311 ESSSSSDSS 320

RESULT 6

US-09-991-262-40
Sequence 40, Application US/09991262
GENERAL INFORMATION:
APPLICANT: Christlan, P. D., Gordon, K. H.J., Hanzlik, T. N.
TITLE OF INVENTION: Insect Viruses and Their Uses in
Protecting Plants
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Flehr Hobdach Test Albritton & Herbert LLP
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/991,262
FILING DATE: 20-NOV-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/234,238
FILING DATE: 20-JAN-1999
APPLICATION NUMBER: US 08/485,355
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/440,522
FILING DATE: 12-MAY-1995
APPLICATION NUMBER: US 08/089,372
FILING DATE: 08-JUL-1993
APPLICATION NUMBER: AU PL4081/92
FILING DATE: 14-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-58631-3/RFT/DSS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-991-262-40

Query Match 1.5%; Score 10; DB 5; Length 1704;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 SSSSSSDSTS 626

Db 1452 SSSSSSDSTS 1461

Search completed: June 13, 2002, 08:49:37
Job time: 228 sec

08/961083

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:40:34 ; Search time 34.61 Seconds
(without alignments)
2137.392 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484
Sequence: 1 KIDNKNQIADLGSESRVN.....TQSNTPDQGNQNPQAP 666

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3484	100.0	666	19	AAW55063
2	3466	99.5	682	17	AAW04359
3	3466	99.5	719	22	AAU37830
4	2206	63.4	420	20	AAU56106
5	1751	50.3	778	22	AAU35135
6	1621	46.5	320	16	AAU70153
7	1361	39.1	462	22	AAU03646
8	817.5	23.5	727	22	AAU34286
9	817.5	23.5	727	22	AAU37184
10	640	18.4	805	21	AAU75603
11	638	18.3	805	21	AAU75602

12	636	18.3	731	19	AAW44849	S. pneumoniae peni
13	635.5	18.2	805	21	AAU75601	Neisseria gonorrhoe
14	625	17.9	660	18	AAW5545	H. pylori ORF 06ep
15	625	17.9	660	18	AAW55576	H. pylori ORF 06ep
16	608.5	17.5	596	20	AAW89836	Protein encoded by
17	586	16.8	812	22	AAU78604	Lawsonia intracell
18	567	16.3	828	17	AAW04357	E. coli penicillin
19	555	15.9	850	17	AAW86955	E. coli PBP 1A tra
20	552	15.8	774	22	AAU36453	Pseudomonas aerugi
21	539.5	15.5	536	20	AAW89885	Antigen 2 from clu
22	529.5	15.2	821	21	AAU81757	Streptococcus pneu
23	526	15.1	823	17	AAW04358	E. coli penicillin
24	526	15.1	844	22	AAU34446	E. coli cellular p
25	525.5	15.1	781	22	AAU35684	Haemophilus influe
26	519.5	14.9	846	22	AAU38311	Salmonella typhi c
27	514	14.8	844	17	AAW86953	E. coli PBP 1B tra
28	512	14.7	844	17	AAW86953	E. coli PBP 1B tra
29	508	14.6	833	22	AAU55400	Propionibacterium
30	489	14.0	836	17	AAW86954	E. coli PBP 1B tra
31	457	13.1	447	18	AAW20510	H. pylori inner me
32	457	13.1	799	22	AAU49462	Propionibacterium
33	450.5	12.9	770	22	AAU98353	Escherichia coli p
34	441.5	12.7	795	22	AAU80056	C glutamicum prote
35	437.5	12.6	678	22	AAU81107	Mycobacterium tube
36	422	12.1	810	22	AAU93052	C glutamicum prote
37	419	12.0	810	22	AAU81232	Mycobacterium tube
38	377	10.8	344	18	AAW55362	H. pylori ORF 14p
39	377	10.8	344	18	AAW20917	H. pylori cell env
40	370.5	10.6	532	17	AAW86957	E. coli penicillin
41	370.5	10.6	553	17	AAW86956	E. coli penicillin
42	341.5	9.8	231	20	AAW50230	R. eutropha Mgt pa
43	288.5	8.3	269	19	AAW61395	Staphylococcus aur
44	287	8.2	338	22	ABG24982	Novel human diagno
45	280.5	8.1	269	22	AAU81893	S. epidermidis ope

ALIGNMENTS

RESULT 1	
AAW55063	standard; Protein: 666 AA.
XX	
AC	AAW55063;
XX	
DF	02-OCT-1998 (first entry)
XX	
DE	Streptococcus pneumoniae SP001 protein.
XX	
KW	Streptococcus pneumoniae: antigen; vaccine; infection; diagnosis;
KW	detection; pneumonia; otitis media; meningitis.
XX	
OS	Streptococcus pneumoniae.
XX	
PN	MO9818930-A2.
XX	
PD	07-MAY-1998.
XX	
PF	30-OCT-1997; 97WO-US19422.
XX	
PR	31-OCT-1996; 96US-0029960.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Choi GH, Hromockyj A, Johnson LS, Kunsch CA;
XX	
DR	WPI: 1998-272224/24.
XX	
DR	N-PSDB; AAV27323.
XX	
PT	Nucleic acid encoding antigenic peptide(s) from Streptococcus
PT	pneumoniae - or their epitope-containing fragments, useful in
XX	protective or therapeutic vaccines, and for diagnosis

PS Claim 11; Page 48; 118pp; English.

CC The present sequence represents a protein from *Streptococcus pneumoniae*.
 CC The nucleic acid sequence encoding the *Streptococcus pneumoniae* protein
 CC can be useful in vaccines for inducing protective antibodies against
 CC *Streptococcus pneumoniae*, for treatment or prevention of infection e.g.
 CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
 CC are used to detect *Streptococcus pneumoniae* infection (by usual hybridisation or
 CC amplification methods), also for isolating *Streptococcus pneumoniae* or their
 CC allelic variants. The protein can be used similarly to detect specific
 CC antibodies in standard immunoassays, especially for diagnosing or
 CC monitoring infections. Antibodies which bind the protein are used to
 CC detect corresponding antigens, to purify the protein and for passive
 CC immunisation (optionally coupled to a toxin). Vaccines are administered,
 CC e.g. by injection, orally or through the skin, typically at 0.01-1000
 CC (especially 10-300) μ g/ml per dose.

CC Sequence 666 AA:

Query Match 100.0%; Score 3484; DB 19; Length 666;
 Best Local Similarity 100.0%; Pred. No. 1,1e-234;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLIGSERRNAQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLEARN 60
 DB 1 KIYDNKNQIADLIGSERRNAQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLEARN 60
 OY 61 LOSNSLOGSGSTLTQOLIKITLFTSTSDQTSRKQAQEAALATOLEKATKQELITYYINK 120
 DB 61 LQSNLSLGGSLTGLTGLIKITLFTSTSDQTSRKQAQEAALATOLEKATKQELITYYINK 120
 OY 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNODPYSHPEAADRRNLVY 180
 DB 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNODPYSHPEAADRRNLVY 180
 OY 181 SEMKNOGYISABQYKAVNTPTDGLQSLKASNPAYMDNKLKVINOVEEETGYNLLT 240
 DB 181 SEMKNOGYISABQYKAVNTPTDGLQSLKASNPAYMDNKLKVINOVEEETGYNLLT 240
 OY 241 TGMADVYNTNDQEAOKHLMQYNTMDEVVAAPDDELQVASTIVVYSNKKVIAQAGARRQSSN 300
 DB 241 TGMADVYNTNDQEAOKHLMQYNTMDEVVAAPDDELQVASTIVVYSNKKVIAQAGARRQSSN 300
 OY 301 VEFGLIQVAETNRDMGSTMKPTIDYAPALEYGYDSTATIVHDEPNYPGTNPVYNNMDR 360
 DB 301 VEFGLIQVAETNRDMGSTMKPTIDYAPALEYGYDSTATIVHDEPNYPGTNPVYNNMDR 360
 OY 361 GFYGNITLQYALQOQRNPVAVETLNNKVLNRAKTFINGLGIDYPSIHYSNNAISSNTTESD 420
 DB 361 GFYGNITLQYALQOQRNPVAVETLNNKVLNRAKTFINGLGIDYPSIHYSNNAISSNTTESD 420
 OY 421 KRYGASSEKMAAAYAFAGNGTYVYKPMYTHKVFVSDGSEKESNVTGRAMKEETAYMMMD 480
 DB 421 KRYGASSEKMAAAYAFAGNGTYVYKPMYTHKVFVSDGSEKESNVTGRAMKEETAYMMMD 480
 OY 481 NAKTITLYTGGRNAYLAWLPQAGKTGTSNYTDEIEENHAKTSQFVAPDCLFAGYTRKRYGM 540
 DB 481 NAKTITLYTGGRNAYLAWLPQAGKTGTSNYTDEIEENHAKTSQFVAPDCLFAGYTRKRYGM 540
 OY 541 AVMTGYSNRLPLVNGSLTVAAYKVRSMNTVYSEGSNPEDMNIPEGLYRNGEVFNKGAR 600
 DB 541 AVMTGYSNRLPLVNGSLTVAAYKVRSMNTVYSEGSNPEDMNIPEGLYRNGEVFNKGAR 600
 OY 601 STWNBPAPQAPSTESSSSSSSSSTPSTNNSTTTTPNNNTTQOQSNTPPDQONON 660
 DB 601 STWNBPAPQAPSTESSSSSSSSSTPSTNNSTTTTPNNNTTQOQSNTPPDQONON 660
 OY 661 POPAP 666
 DB 661 POPAP 666

RESULT 2

ID AAW04359 standard; Protein; 682 AA.

AC AAW04359;

DT 03-DEC-1996 (first entry)

DE S. pneumoniae penicillin binding protein 1A soluble variant.

KM Penicillin binding protein; PBP 1A; bifunctional protein;

KW transglycosylase; transpeptidase; identification; assay; inhibitor;

KW antibiotic resistant; bacteria; soluble variant; protein structure;

OS Streptococcus pneumoniae.

PN GB2290792-A.

PD 10-JAN-1996.

PE 29-JUN-1995; 95GB-0013306.

PR 24-NOV-1994; 94SE-0004072.

PR 01-JUL-1994; 94TN-0000580.

PA (ASTR) ASTRA AB.

PI Balanesh TS, Town CM;

DR WPI; 1996-042332/05.

DR N-PSDB; AAT08027.

PT Sol. derivs. of bifunctional penicillin binding protein (BPBP) -

PT opt. lack transglycosylase activity, useful to identify and assay

PT for antibodies or cpds. which bind BPBPs

PS Claim 2; Pages 65-68; 108pp; English.

XX The present sequence is a soluble variant of the S. pneumoniae

CC penicillin binding protein (PBP) 1A, which lacks the 1st 38 amino

CC acid residues of the wild type protein. Wild type PBP is a

CC bifunctional protein, which binds the cell membrane when expressed

CC in a bacterial cell, having transglycosylase and transpeptidase

CC activities. The variant protein (NCIMB 40665) in conjunction with

CC a labelled anti-bifunctional PBP monoclonal antibody, can be used

CC to identify and assay for cpds. which bind bifunctional PBP. Such

CC cpds., as inhibitors of bifunctional PBP have a potential use in

CC therapeutic cpds. which inhibit the growth of antibiotic resistant

CC bacteria. The soluble variant may also be used in x-ray

CC crystallography.

XX Sequence 682 AA:

Query Match 99.5%; Score 3466; DB 17; Length 682;
 Best Local Similarity 99.4%; Pred. No. 2e-233;
 Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 KIYDNKNQIADLIGSERRNAQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLEARN 60
 DB 1 KIYDNKNQIADLIGSERRNAQAANDIPTDLVKAIVSIEDHREFDHRGIDTIRILGAFLEARN 60
 OY 61 LOSNSLOGSGSTLTQOLIKITLFTSTSDQTSRKQAQEAALATOLEKATKQELITYYINK 120
 DB 61 LQSNLSLGGSLTGLTGLIKITLFTSTSDQTSRKQAQEAALATOLEKATKQELITYYINK 120
 OY 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNODPYSHPEAADRRNLVY 180
 DB 121 VYMSNGNYGMQTAQNYKYKDLNNLSLPOLALLAGMPQAPNODPYSHPEAADRRNLVY 180
 OY 181 SEMKNOGYISABQYKAVNTPTDGLQSLKASNPAYMDNKLKVINOVEEETGYNLLT 240
 DB 181 SEMKNOGYISABQYKAVNTPTDGLQSLKASNPAYMDNKLKVINOVEEETGYNLLT 240

Db 197 semknggyisaeyekavntpitdglqslksasnypaymdnlykevingveetgynlilt 256
QY 241 TGMVYTNVDQEAQKHLMDIYNTDEYVAYPDDDELQVASTIVDVSNGKVIQAOLGARHQSNN 300
Db 257 tgmdivtnvdgeaqkhlwdiynntdeyaypddelqvastivdvsgkviaglgarhqsnn 316
QY 301 VSFGINQAVETNRDWSGTMKPIITDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNNWDR 360
Db 317 vsfginqavetnrdwsgtmkpiitdyapalegyvystativhdepyngtntpyynwdr 376
QY 361 GYFGNITLQYALQOQRNVPVAVETLNKVLNRAKTLFNLGLGIDYPSIHYSNALSSNTTESD 420
Db 377 gyfgnitiqlaylqqrnvpavetlnkvglnraktfnglidiydpshysnalssnttesd 436
QY 421 KKYGASSEKMAAAYAAFAANGGTYYPKPMYIHKVVFSDGSEKFSNVGTRAMKETAYMMTD 480
Db 437 kkygassekmaaaayaaafanggttyypkpmihkvvfsgdsekefsnvgtramkettaymmtid 496
QY 481 MMKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEEIHNIKTQSOFVAPDELFAGYTRKYSM 540
Db 497 mmktvlsygtgrnaylawlpqagktgtsnytdedeiheniktsqfvapdelafagytrkysm 556
QY 541 AVWTGYSNRLTPLVGNGLTVAAKVYRSMYTLSEGSNPEDWNIPGELYRNGEFVKNGAR 600
Db 557 avwtgysnrltplvnglntvaakvyrsmttylsegnsnpedwnipeglyrngefvfkngar 616
QY 601 STWNSPAPQPPSTESSSSSSSTSSQSSSTPSTNNSTTTNPNNNTQOSNTTTPDQONON 660
Db 617 stwspapqppstesssssdssstssqsssttpstnnstttnpnnntqosnttppdqonqn 676
QY 661 PQAQAP 666
Db 677 pqapqp 682
RESULT 3
AAU37830
ID AAU37830 standard; Protein: 719 AA.
XX AC AAU37830;
XX XX
DT 14-FEB-2002 (first entry)
XX DE Streptococcus pneumoniae cellular proliferation protein #259.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Streptococcus pneumoniae.
XX FN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207727P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX XX
XX PI Haselbeck R, Ohlson KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yanamoto RT, Xu HH;
XX DR WPI: 2001-611495/70.
XX DR N-PSDB; AAS55689.
XX PT New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -
XX Example 3; Seq ID No 13423; 511pp; English.
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 719 AA;
Query Match 99.5%; Score 3466; DB 22; Length 719;
Best Local Similarity 99.4%; Pred. No. 2.1e-233;
Matches 662; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 KIYDNKNQIADLGSERRVNAOANDIPTDLVKAIVSIEDHRRFDHGRIDITRILGAFLRN 60
Db 54 kiydhnkqliadlgserrvnaoandiptdlvkaivsiedhrfridhrgiditirilgaflrn 113
QY 61 LOSNSLQGGSTITQOLIKUTYFSTSDQTSIRKAQEAWLAIQLEOKATKQBELTYIINK 120
Db 114 lqsnslqggstaltqliklityfststsdqtsirkaqeaawlaiqlqekatkqbeltyiink 173
QY 121 VYMSNGNYGMQTAQNYGYKDLNLSLPOLALLAGMPQAPNOYDYPHPEAQRNLVL 180
Db 174 vymnsngnygmtaagnygykdlnlsipqlallagmpqapndydpshpeaqrnlvl 233
QY 181 SEMKNGGYISAEQYEKAVNTPIITDGLQSLKSASNPAYMDNLYKEVINOVEBETGYNLT 240
Db 234 semknggyisaeyekavntpitdglqslksasnpaymdnlykevingveetgynlilt 293
QY 241 TGMVYTNVDQEAQKHLMDIYNTDEYVAYPDDDELQVASTIVDVSNGKVIQAOLGARHQSNN 300
Db 294 tgmdivtnvdgeaqkhlwdiynntdeyaypddelqvastivdvsgkviaglgarhqsnn 353
QY 301 VSFGINQAVETNRDWSGTMKPIITDYAPALEYGVYDSTATIVHDEPNYPGTNTPYNNWDR 360
Db 354 vsfginqavetnrdwsgtmkpiitdyapalegyvystativhdepyngtntpyynwdr 413
QY 361 GYFGNITLQYALQOQRNVPVAVETLNKVLNRAKTLFNLGLGIDYPSIHYSNALSSNTTESD 420
Db 414 gyfgnitiqlaylqqrnvpavetlnkvglnraktfnglidiydpshysnalssnttesd 473
QY 421 KKYGASSEKMAAAYAAFAANGGTYYPKPMYIHKVVFSDGSEKFSNVGTRAMKETAYMMTD 480
Db 474 kkygassekmaaaayaaafanggttyypkpmihkvvfsgdsekefsnvgtramkettaymmtid 533
QY 481 MMKTVLTYGTGRNAYLAWLPQAGKTGTSNYTDEEIHNIKTQSOFVAPDELFAGYTRKYSM 540
Db 534 mmktvlsygtgrnaylawlpqagktgtsnytdedeiheniktsqfvapdelafagytrkysm 593
QY 541 AVWTGYSNRLTPLVGNGLTVAAKVYRSMYTLSEGSNPEDWNIPGELYRNGEFVKNGAR 600
Db 594 avwtgysnrltplvnglntvaakvyrsmttylsegnsnpedwnipeglyrngefvfkngar 653
QY 601 STWNSPAPQPPSTESSSSSSSTSSQSSSTPSTNNSTTTNPNNNTQOSNTTTPDQONON 660
Db 601 stwspapqppstesssssdssstssqsssttpstnnstttnpnnntqosnttppdqonqn 660

Db 654 stvsspapqpppsssssdssstgsssttpstnnstttppnnnttqgsnttpdqgnqn 713

QY 661 POPAQP 666
|||||

Db 714 pqpapq 719

RESULT 4
 AAY56106
 ID AAY56106 standard; Protein; 420 AA.
 AC AAY56106;
 XX
 XX 07-FEB-2000 (first entry)
 XX Streptococcus pneumoniae pbp1A TER isolate a) protein sequence.
 XX
 KW Streptococcus pneumoniae; penicillin binding protein; pbp2B; pbp1A;
 KW transpeptidase encoding region; TER; antibiotic resistance; diagnosis;
 KW detection; identification; pneumococcal meningitis.
 KW
 XX
 OS Streptococcus pneumoniae.
 XX
 XX ZA9807024-A.
 PN XX
 PD 28-APR-1999.
 PD XX
 PF 05-AUG-1998; 98ZA-0007024.
 PF XX
 XX 01-AUG-1997; 97ZA-0006886.
 XX PR
 XX (SAME-) SOUTH AFRICAN INST MEDICAL RES.
 XX PA (UYWT-) UNIV WITWATERSRAND.
 XX PA (MEDI-) MEDICAL RES COUNCIL.
 XX PA
 XX Klugman KP, Smith AM, Du Plessis M;
 XX PI WPI; 1999-601770/51.
 XX DR N-PSDB; AA235939.
 DR XX
 XX
 PT Polymerase chain reaction assays for detecting Streptococcus pneumonia
 PT useful for the diagnosis of pneumococcal meningitis
 PT
 PS Claim 11; Fig 4; 63pp; English.
 PS
 XX
 CC A polymerase chain reaction (PCR) assays have been developed for
 CC detecting an antibiotic resistant strain of Streptococcus pneumoniae
 CC using primers based on the penicillin binding protein 2B (pbp2B) gene
 CC and the pbp1A gene. The products and methods can be used for detecting
 CC S. pneumoniae, particularly antibiotic-resistant strains. They can be
 CC used for simultaneously diagnosing pneumococcal meningitis and
 CC identifying any antibiotic-resistant S. pneumoniae strains in a sample
 CC The methods can be used for detecting S. pneumoniae strains resistant
 CC to antibiotics, e.g. beta-lactam antibiotics, preferably penicillin.
 CC The assays can be adapted to detect other pathogens causing meningitis
 CC The assays can be used to detect an antibiotic resistant strain of
 CC S. pneumoniae with a minimum inhibitory concentration (MIC) of
 CC 0.25-1 micro g/ml where the PCR products detected are a 1043 bp and
 CC a 224 bp product. The present sequence represents a Streptococcus
 CC pneumoniae pbp1A transpeptidase encoding region (TER) isolate protein
 CC sequence from the present invention.
 CC
 XX Sequence 420 AA;
 SQ

Query Match 63.4%; Score 2208; DB 20; Length 420;
 Best Local Similarity 99.0%; Pred. No. 6.6e-146;
 Matches 416; Conservative 3; Mismatches 1; Indels 0; Gaps

QY 192 EQYEKAVNTPITDGLQSLKSNYPAYMDNVLKEVINGVEETGYNLLTTGMDVYTNVDQ 251
 ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 eggykavntpITdGLqslksasnyPAYmdnYLkevinqveetgynllttgmDvYtnvdq 60

CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX

Sequence 778 AA;
Query Match 50.3%; Score 1751; DB 22; Length 778;
Best Local Similarity 51.4%; Pred. No. 1.3e-113;
Matches 347; Conservative 125; Mismatches 193; Indels 10; Gaps 3;
QY 1 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHRFFDHGIDTIRILGAFNRN 60
Db 81 kiydneifeidgaekrelidqndvpqllkdaivsvedrfrfkyhigvpiriigsalsn 140
QY 61 LQSNLSGGSTLTQOLIKLTYFSTSTSDQTSRKAQEAWLAIQLEQKATKQEILTYINK 120
Db 141 vknglgggstltqllklyfsfksesdqltkrkaqeamavrlerekseelityink 200
QY 121 VYMSNGYGMQTAQNYGKDLNLSLPQALLAGMPQAPNQDPYSHPEAAQDRNLVL 180
Db 201 vyangfygmataenygghelddipqtallagmpqapnsydytkpdtakerrdvvl 260
QY 181 SEMKNOCYISAEQYKAVNTPIITDGLQSLKSASNYPAYMDNLYKEVINQVEEETGYNLLT 240
Db 261 ytnydnkkiskaeakakatpideglvplkasddnrkvdyvkevinevaktgknvyt 320
QY 241 TGMDEVYNDVDOEAQKHLWDIYNTDEYVAYPDDELQVASTIVDSVNGKVIQAQLGARHOSN 300
Db 321 dgldiytnldmnaqkqlydvnsdqyvaqfddkmqvastvidvasgqvraqiggrhpd 380
QY 301 VSFGINQAVETNRDWSGTMKPIITDYAPALEYGVYDSTATIVHDEPYNYPGTNTFVYNWDR 360
Db 381 vqlgnnlavntqrdvgstvkpmdygaipenlny-stgrlmvdktpkypgtddvfnsl 439
QY 361 GYFCNITLQVALQSRNPVAVETLNKVLNRAKTFLLGLGIDYPSIHYNSAISNTTRES 420
Db 440 tygvtmrainsrnttavqtfdevrgkenimpfkgldgdyknleasnaissntsdvd 499
QY 421 -KKYGASSEKMAAAYAFANGGTYYPKYIHKVVFSDGSEKFSNVGTRAMKETAYMMT 479
Db 500 gkygisslklaaayafanggtyynkpyykvvfnvndgtsvdyqpdgkramkdstaymnt 559
QY 480 DMKTVLTCTGNAYLAWLPQAGKCTGSNTYDDEIENHIKTSOFVAPDELFACTRYKS 539
Db 560 dmladvnggctgngaipgliaakqtsnytdedlarmttekgiapdstfvgytthya 619
QY 540 MAYVTGYSNRLTPLVGNGLTVAARVYRSMYTLSEGSNPEDWNPEGLYRNGEYFVKNGA 599
Db 620 vsvwtgynrntpiyqeyygiasdvreinsylsqvnsndvqpdsvrvrgnelyykda 679
QY 600 RSTWN-----SPAPQPPSTESSSSSDSTTSQSSSTPTPSNNSTTTNNNNNTQOOSN 651
Db 680 yevpnyqlpsttssapqpsstssstkeasessssssseapssseapstqpass 739
QY 652 TTPDQONQNPQAP 666
Db 740 ssaeqpatseqppep 754

RESULT 6
AAR70153
ID AAR70153 standard; Protein: 320 AA.

XX AAR70153;
XX AC 14-FEB-1996 (first entry)
XX DT Streptococcus pneumoniae strain SPRU42 Exp2.
XX DE Exp2; export protein; bplA; plpA; exp1; exp3; pad1;
XX KW virulence determinant; permease like protein;
XX KW penicillin binding protein 1A; pyruvate oxidase; regulatory element;
XX KW acellular vaccine; antibody.
XX OS Streptococcus pneumoniae.
XX PN WO9506732-A2.
XX PN 09-MAR-1995.
XX PD 01-SEP-1994; 94WO-US09942.
XX PF 01-SEP-1993; 93US-0116541.
XX PR 18-MAY-1994; 94US-0245511.
XX XX (UYRQ) UNIV ROCKEFELLER.
XX Masure HR, Pearce BJ, Tuomanen E;
XX WIPI: 1995-115448/15.
XX DR N-PSDB; AAQ83241.
XX Novel gene fragments encoding specific bacterial exported proteins
XX - specifically of S. pneumoniae, useful as vaccines
XX Claim 35; Page 88-9; 168pp; English.
XX This sequence represents exp2. The DNA encoding this sequence is
XX identical to that for ponA which encodes penicillin-binding protein 1A
XX (bplA). This sequence is involved in adhesion of bacteria to target
XX cells. This sequence is an exported protein of S.pneumoniae. Export
XX proteins are the proteins in pathogenic bacteria that are virulence
XX determinants. Other export proteins include plpA (see AAR70152), exp1,
XX exp3, and pad1 (encoded by the sequence shown in AAQ83259). This
XX sequence can be inserted into an expression vector (preferably a
XX bacterial expression vector) to provide for high levels of expression of
XX the protein. The protein can then be used in the production of an
XX acellular vaccine. These vaccines are used to provide protection from
XX Gram positive bacterial infection. Antibodies against export proteins
XX can be used for diagnosis of infection and in passive immune therapy.
XX SQ Sequence 320 AA;
Query Match 46.5%; Score 1621; DB 16; Length 320;
Best Local Similarity 99.7%; Pred. No. 4.3e-105;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 KIYDNKNOLIADLGSERRVNAQANDIPTDLVKAIVSIEDHRFFDHGIDTIRILGAFNRN 60
Db 5 kiydnknoliadlgserrvnaqandiptdlvkaivsiedhrffdhgdtirilgafnrn 64
QY 61 LQSNLSGGSTLTQOLIKLTYFSTSTSDQTSRKAQEAWLAIQLEQKATKQEILTYINK 120
Db 65 lqsnlsggstltqllklyfstsdsdtlsrkaqeaawlaiqlqekatkqeilkyink 124
QY 121 VYMSNGYGMQTAQNYGKDLNLSLPQALLAGMPQAPNQDPYSHPEAAQDRNLVL 180
Db 125 vymngnygmqttaaqnyygdlnlslpqlallagmpqapnqdpyshepaqdrnlvl 184
QY 181 SEMKNOCYISAEQYKAVNTPIITDGLQSLKSASNYPAYMDNLYKEVINQVEEETGYNLLT 240
Db 185 semknogyisaeqyekavntpitdglqslksasnypaymdnlykevinqveeetgynllt 244
QY 241 TGMDEVYNDVDOEAQKHLWDIYNTDEYVAYPDDELQVASTIVDSVNGKVIQAQLGARHOSN 300

|||||
245 tgmdivydvdeaqkhlwdyntdeyaypddelqvastivdvsnkgkviaglgarhgsn 304
QY 301 VSFQINQAVETNRDWG 316
Db 305 vsfgingqvavetnrdwg 320
RESULT 7
AAU03646
ID AAU03646 standard; Protein; 462 AA.
XX
AC AAU03646;
DT 12-SEP-2001 (first entry)
XX
DE Group B Streptococcus antigenic protein, ID-122.
XX
KW Group B Streptococcus; encapsulated bacterium; therapeutic; sepsis;
KW meningitis; neonate; antigenic; vaccine; infection; genital tract;
KW capsid polysaccharide vaccination.
XX
OS Streptococcus agalactiae.
XX
WO200132882-A2.
XX
PD 10-MAY-2001.
XX
PF 07-SEP-2000; 2000WO-GB03437.
XX
PR 07-SEP-1999; 99GB-0021125.
XX
PA (MICR-) MICROBIAL TECHNIQS LTD.
XX
PI Le Page RWF, Wells JM, Hanniffy SB;
XX
DR WPI; 2001-316444/33.
DR N-PSDB; AAS07063.
XX
XX New polypeptides derived from Streptococcus agalactiae are useful to
PT provide detection of, and vaccination against, Group B Streptococcus
PT infections, particularly to prevent infection in neonates -
XX
PS Claim 1; Fig 1; 178pp; English.
XX
CC AAU03601-AAU03722 represent Group B Streptococcus (Streptococcus
CC agalactiae) amino acid sequences of the invention.. S. agalactiae is an
CC encapsulated bacterium which is a major pathogen of humans causing sepsis
CC and meningitis in neonates as well as adults. The S. agalactiae antigenic
CC polypeptides are used to vaccinate against Group B Streptococcus
CC infections, particularly to prevent infection in new born children
CC arising from the maternal genital tract. An immunogenic composition is
CC useful in the preparation of a medicament for the treatment or
CC prophylaxis of Group B Streptococcus infection. The invention does not
CC have the disadvantages of varied response rate associated with prior art
CC capsid polysaccharide vaccination against Group B Streptococcus.
XX
SQ Sequence 462 AA;

Query Match 39.1%; Score 1361; DB 22; Length 462;
Best Local Similarity 62.0%; Pred. No. 1e-86;
Matches 251; Conservative 73; Mismatches 81; Indels 0; Gaps 0;
QY 2 IYDNKNLIADLGSERRVNAQNDIPTDLKAIYSDHEDHRRGIDTIRILGAFRLNL 61
Db 58 vvdgnklladlgseksvsadipnlvnaitscedkfrffkhrvgdylrilgaahnl 117
QY 62 QSNLSQGGSTLQILKILTYFSTSDQTSRKAQAWLAQLQKATKQILYIYINKV 121
Db 118 vsntqggstldqllklayfstnksqtkrksqgwlalqmerkykeilfyinkv 177
QY 122 YMSNGNYGMTAAQNYGKDLNLSLPQALLAGMPQAPNOYDPSHPEAAQDRRLVLS 181

|||||
178 ymngnygmrttaksyfgkdkelsiaqlallagipdaptqdydknpesaaqrntvlq 237
QY 182 EMKNQGYISAEQYEKAVNTPTDGLQSLKASNSPAYMDNYLKEVINQVEEETGYNLLTT 241
Db 238 qmyqdkniskkeydqavatpvtgkkelkqkstydkymdnylkqvisevkgktgkdifta 297
QY 242 GMDVYTNVDQEAQKHLWDIYNTDEYVAYPDDELQVASTIVDVSNKGKVIAGLGRHQSNNV 301
Db 298 glkvytnintdaqkqlydiynsdtyiaypnnelqastimdatngkviaglggrhgheni 357
QY 302 SFGINQAVETNRDWGSTMKPTIDYAPALEYGVYDSTATIVHDEPNYPGNTPTVYNDRG 361
Db 358 sfgtnqsvlttdrgwgtmkpisyapaldsgvynustggsindsvyypgtstqlydwdrq 417
QY 362 YFGNITLQYALQOSRNPVAVETLNKVLNRAKTEFLNGLGDYPSI 406
Db 418 ymgwmsmqtaiqgrnvpavraleaagldaeksflekglyypem 462
RESULT 8
AAU034286
ID AAU034286 standard; Protein; 727 AA.
XX
AC AAU034286;
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #562.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-611495/70.
DR N-PSDB; AAS52145.
XX
PT New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 5782; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 727 AA;

Query Match 23.5%; Score 817.5; DB 22; Length 727;
Best Local Similarity 32.1%; Pred. NO. 1.8e-48;
Matches 222; Conservative 128; Mismatches 261; Indels 81; Gaps 22;

QY 1 KIYDNKNQIADL-GSERRVNAQANDIPTDLVKAIVSIEDHRRFFDHRGIDTIRILGAFRL 59
DB 76 kiydngelvktdlmgqrhehvnldkvpksmkdaviatednrfyehgaldykrifgaigk 135
QY 60 NLOSN-SLOGGSTLTQQLIKLYFTSTSDQTSIRKAQAWLAIQLEQKATQOEILTYII 118
DB 136 nltgfgsegastltqgvvkdafsl---qhksigrkaeaylsyrleqeyskddifqyvl 192
QY 119 NKVYMSNGYGMOTRAQNYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPEAQRRL 178
DB 193 nkiiysdgvgtgikaakyfynkldlnlaeeaylaglpqvpnnyniylhphkkaedrnt 252
QY 179 VLSEMKNGYISAEQYKA-----VN-TPITDGLQLSKASNYPAYMDNYLKEVIN 228
DB 253 vlylmhykridkwdedakkidkanlvnrcteerqnditqdsyynsvnfkselmm 312
QY 229 Q--VEEETGYNLLTGMVDYVNDQEAQKHLWDIYNTDEYVAYPDEQLQVASTIVDVSNG 286
DB 313 mkafkdenlgnvlqsgiklyntmdkdvktlgn--dvdngsfyknkdgqvgatildsktg 370
QY 287 KVIAQLGARHOSNNVFGINQAVETNRDWMGTMKPIDYAPALEYGVVDSTATIVHDEPY 346
DB 371 glvaisggrdfkdvvn--rnqatdph-ptgssalkpflagypaiekmkwatnhaigdesy 427
QY 347 NYPGNTPTVYNDRGYFNGITLQYALQOSRNVPAVETLNKV---GLNRAKTFNLGLGD 402
DB 428 qvdgst--frnydtkshgtcvsllydairgsfnlpalkawgsvkqngnagdapkkfaaklgln 485
QY 403 YP-SIHYNIAISNPTSDKKYIGASERKMAAAYAFANGTYTKPMYIHKVVFSDGSEKE 461
DB 486 yegdipgpsevlgsase-----fsgtqlasafaiaanggtynnahsiqkvvtrdgetie 539
QY 462 FSNVGTAMKETAYMMTDMK-TVLVTGTGRNAYLAMLPOAGKGTGTSNYTDEEIHNIK 520
DB 540 ydhtshkamsdytamlaeamlkgtfkpgysayghvgvngmaktgtygaety----- 594
QY 521 TSQFVAPDEL-----FAGYTRKYSNAVWTGYSNRLTPLVGNGLTVAAKVYRSMWYLSG 575
DB 595 -sqynlpdnakdwingftpqytnsvvmgtsk-----vkqy-----g 631
QY 576 SNP-----EDWNIPEGLYRNGEFVFKNGARSTWNSPAPQPPSTESSSSSDSTSQSSST 631
DB 632 ensfvgshsqeypqflyen--vmskissr---dgedfkrpssvsgsipsinvgsgqdnt 686
QY 632 TPST---NNSTTNPNNNTQQSNTPDQONQN 660
DB 687 tnrtshggsdtsanssgtaqsnnttrsggsrn 718

RESULT 9

AAU37184
ID AAU37184 standard; Protein; 727 AA.
AC AAU37184;
XX
XX
DT 14-FEB-2002 (first entry)
XX

DE Staphylococcus aureus cellular proliferation protein #1354.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI; 2001-6111495/70.
DR N-PSDB; AAS55043.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 12777; 511pp; English.
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 727 AA;

Query Match 23.5%; Score 817.5; DB 22; Length 727;
Best Local Similarity 32.1%; Pred. No. 1.8e-48;
Matches 222; Conservative 128; Mismatches 261; Indels 81; Gaps 22;
QY 1 KIYDNKNQIADL-GSERRVNAQANDIPTDLVKAIVSIEDHRRFFDHRGIDTIRILGAFRL 59
DB 76 kiydngelvktdlmgqrhehvnldkvpksmkdaviatednrfyehgaldykrifgaigk 135
QY 60 NLOSN-SLOGGSTLTQQLIKLYFTSTSDQTSIRKAQAWLAIQLEQKATQOEILTYII 118
DB 136 nltgfgsegastltqgvvkdafsl---qhksigrkaeaylsyrleqeyskddifqyvl 192
QY 119 NKVYMSNGYGMOTRAQNYGKDLNNLSLPQALLAGMPQAPNOYDPYSHPEAQRRL 178
DB 193 nkiiysdgvgtgikaakyfynkldlnlaeeaylaglpqvpnnyniylhphkkaedrnt 252

QY	179	VLEMRKNOGVISAEQYEKA-----VN-TPITDGLQSLKASNPAYMDNYLKEVIN	229
DB	253	VLYIMHYHKRITDKQWEDAKIKLkanlvrtppeerqnltndqdsyasyvafvkselmm	312
QY	229	Q--VEETGYNLLTGMGVYTVNDVQEAQKHLMDIYNTDEYVAYPDEELQVASTIVDVNSG	286
DB	313	nkafkdnglgnvlqsgiklyltmmdkdvqtklgn--dvdngsfyknkdgqvgatllidsktg	370
QY	287	KVIAQLGARHOSQSNVSGFINGAVETNRDQSGTMMKPTTDYAPALEYGVYDSTATIVHDPEY	346
DB	371	glvaisggrdfkdvnn--rnqatdph-ptgsslkpfaygpaiaemkwtatnhaicqdesy	427
QY	347	NYPGTNPVYVNDRGVYFGNITLQYALQOGRNVPVAVETLNKV---GLNRAKTFLLNGLGID	402
DB	428	qvdst--frnydtkshgtvsiydalrsgfnipalkawgsvkqngagndapkkfaaklgl	485
QY	403	YP-SIHYSNAISSNTTSEDDKKYCASSEKMAAAYAAFANGGTYYKPMYIHKVVFPDGSSEKE	461
DB	486	yegdlpsevlgsase-----fspqtqlasafaalangtgyynnahsqkvvttrdgetie	539
QY	462	FSNVGTRAKETAYMMTDMKK-TVLTYTGTRNVALWLPQAGKTGTSNYTDEETENHIK	520
DB	540	ydhtsnkamsdytaymlaemlkgtfkpygsahyghvsgvnmqaktgtgygaety-----	594
QY	521	TSQFVAPDEL-----FAGYTRYKSYMAVWTGYSNRLTPLVGNGLTVAAKYVYRSMYTYLSEG	575
DB	595	-sgynlpdnaakdwingftpgytmvsvmngfsk-----vkgqy-----g	631
QY	576	SNP-----EDWNIPEGLYRNCEGFVKNGARSTWNSPAPQPPSTESSSSSSSSSTSQSSST	631
DB	632	ensfvghsqgeypqfilyen--vmsklssr---dgedfkrpssvsgslpsinvsqsgdnt	686
QY	632	TPST---NSTTTNPNNNTQQSNTTTPDOQNQN	660
DB	687	tnrsthgdsdtsaansgtagssnnntrsqsrn	718
RESULT 10			
AAV75603			
ID	AAV75603 standard; Protein; 805 AA.		
XX			
AC	AAV75603;		
XX			
DT	21-MAR-2000 (first entry)		
XX			
DE	Neisseria meningitidis ORF 791 protein sequence SEQ ID NO:2680.		
XX			
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;		
KW	antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;		
KW	antibacterial; gene therapy.		
XX			
OS	Neisseria meningitidis.		
XX			
PN	W0957280-A2.		
PD	11-NOV-1999.		
XX			
PF	30-APR-1999; 99WO-US09346.		
XX			
PR	01-MAY-1998; 98US-0083758.		
PR	31-JUL-1998; 98US-0094869.		
PR	02-SEP-1998; 98US-0098994.		
PR	02-SEP-1998; 98US-0099062.		
PR	09-OCT-1998; 98US-0103749.		
PR	09-OCT-1998; 98US-0103794.		
PR	09-OCT-1998; 98US-0103796.		
PR	25-FEB-1999; 99US-0121528.		
XX			
PA	(CHIR) CHIRON CORP.		
PA	(GENO-) INST GENOMIC RES.		
XX			
PI	Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;		
PI	Petersen J, Piza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;		

[illegible]

trans-peptidation activity; infection; therapy.
Streptococcus pneumoniae.
EP837132-A2.
22-APR-1998.
17-OCT-1997; 97EP-0308288.
17-OCT-1996; 96US-0731716.
(ELIL) LILLY & CO ELI.
Hoskins JA, Jaskunas SR, Norris FH, Rostock PK, Rostock PR;
Zhao G;
WPI; 1998-219114/20.
N-PSDB; AAV19373.
Recombinant Streptococcus pneumoniae penicillin-binding protein -
useful in screening assays for antibacterial agents
Claim 1; Page 19-21; 28pp; English.
This sequence is the Streptococcus pneumoniae penicillin-binding
protein, designated PBP-NV, of the invention. The protein is useful in
screening assays for compounds that bind to PBP-NV or inhibit the
transglycolase or trans-peptidation activity of PBP-NV. Such compounds
would be useful as antibacterial agents for treating S. pneumoniae
infections.
SQ Sequence 731 AA;
Query Match 18.38; Score 636; DB 19; Length 731;
Best Local Similarity 27.78; Pred. No. 8.5e-36;
Matches 192; Conservative 108; Mismatches 252; Indels 142; Gaps 19;
QY 2 IYDNKNQLIADLGSERRVNAQANDIPTDLKVAIVSIEDHREFDHRGIDTIRILGAFRLNL 61
Db 95 lfdreekeagalsgqkytveltdsklnqlnaviatedrsfykndginygrf---flaiv 151
QY 62 QSNSLQGGSTTQOLIKLYTSTSDQTSIRKQAQEAWLAIQLEQAKFKQBIILTYINKV 121
Db 152 tagrsgggsttltqlaknayls---qdtverkakelflaelskyskeqiltmylnna 208
QY 122 YMSNGNYGMQTAQONYGKDLNNLSLPOLALLAGMPQAPNOYDPYSHPEAQDRRLNVL 181
Db 209 yfgngvvgvedaskkyfvsasevslqdaatlagmlkqepelynpinsvedstnrrdtvlq 268
QY 182 EMKNOGYISAEQYKAVTPIIT---DGLQSLKSASNPAYMDNYLKEVI---NQVEEET 234
Db 269 nmvaagyidknqeteaevadntslhdkyegkisdrypsfydavnveavskynlteee- 327
QY 235 GYNLLTGMVYTNVDQEAQKHLMDIY-NTDEYVAYPDDEL-QVASTIVDVSNKGVIQAL 292
Db 328 ---lvngyrylteidqnyqanmqiventslfpraedgtfaqsgsvalepkgtggrvrv 384
QY 293 GARHQSSNVSF-GINQAVETNRDMGSTMKPIITDYAPALEYGYVDSTATIVHDEPY----- 346
Db 395 gqvadndktgrfnfaytqskrspgtkplvytpaveagwainkqldnhtmgdydykv 444
QY 347 -NYPGTNT-----PVYNWDRGYFGNITLQYALQSSNPVAVETLNKVLNRAKTFNLGLGI 401
Db 445 dnyagiktsrevpmqy-----slaeslnlpavatvndlgvdka----- 482
QY 402 DYPSTHYSNAISSNTTESDKKYGASSEK-----NAAAYAFANGGT 442
Db 483 -----feagekfglnmekvdrvlvgalgsgetpnlqmaayafanegl 527
QY 443 YKPMYIHKVVFSOGSE-KEFSNVGTRAMKETAYMTDMKTVLTGTGRNAYLAWLPQ 501

Db 528 mpeahfisrienasgqviasnksqkrvidksvadmktmmmlgtftngtgiisspadym 587
QY 502 AGKTGTGNTYDEIEENHIKTSQFVAPDELFAGYTRKYSMAYVTGY-----SNRLTPLVGN 556
Db 588 agktgtt-----eavfnpeyts-----dqwvigtpdvvishwlgfpttdenhylagstsn 638
QY 557 GLTVAAKVYRSMYTLSEGSNPEDWNIPEGLYRNGEFYFKNGARSTWNSPAPQPPSTES 616
Db 639 g---aahvfrniantil-----pytpg 657
QY 617 SSSSDSSTSSSTSTPSTNNSTTNNPNNNTQOS 650
Db 658 stftvenaykngiapantkrqvqtdnustqddn 691
RESULT 13
AA75601
ID AAY75601 standard; Protein; 805 AA.
XX
AC AAY75601;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria gonorrhoeae ORF 791 protein sequence SEQ ID NO:2676.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
XX WPI; 2000-062150/05.
DR N-PSDB; AAZ54363.
XX
PT Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
PS Claim 2; Page 1267; 1453pp; English.
XX
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAZ75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ5473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention

Db 165 nqtffghygyvktaslgfyfkkpdkldlklkeitmvalprapsfydptknlefsrand 224

Qy 179 VLSEMKNGYISAEQYKAVN-TPITDGLQSLKASNPAYMDNLYKEVINQVEEETGYN 237

Db 225 llrrlyslgwissnelksalnepivynqtstgniapy-----vdevlklqldlg-- 276

Qy 238 LLTTCMDVYTNVDQEAQKHLWDIY-----NTDEYVAYPDDELQVASTIV 281

Db 277 lktggytkltdldyqrlaleslrfghqkilekiaekpktnasndkdednlnasmivt 336

Qy 282 DVSNGKVIQALGARHSSNVSGFNGQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIV 341

Db 337 ectstgkialalvggidykks---afnratqakrfgsaakpfv-yqlafdnngy--sttski 390

Qy 342 HDEPNYPTGTPVY-----NMDRGYFGNITLQYALQOQSRNPVAVETLNKVG 388

Db 391 pdtarnfengn---ysknsqvphawhpsnytrkfglvtlqlealslnlatinsldqgl 447

Qy 389 LNRKATFUNGIGIDYPSIHYSHNAISSNTTESDKKYGASSEKMAAAYAAFANGGYTKPMY 448

Db 448 fekiyqslsdmgfknlpkdlslvigs-----faispidaaeakyslfsnygtmklpml 499

Qy 449 IHKVVFSDSSEKFSNVGTRAM-KETAYMMTDMKTKVLTGTGRNAYLAWLPOAGKTGT 507

Db 500 iesitnqqnevktftptetkkitkeqafitlsalmdavengtgsarikgleiagktgt 559

Qy 508 SNYTDDEEIEENHIKTSQFVAPDELFGYTRKYSMAVWTGYSNRLTPLVGNGLT---VAARKV 564

Db 603 ysfmrnilaiepslkrfdvpkglrk--eivdkipyyspnsitp-tpkkttds 654

RESULT 15

AAW55576

ID AAW55576 standard; Protein: 660 AA.

XX AC AAW55576;

XX AC AAW55576;

DT 07-JUL-1998 (first entry)

DE H. pylori ORF 06ep30223_4698838_f2_55 protein.

XX KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;

XX KW identification; binding compound; bacteria; life cycle; activator;

XX KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.

XX OS Helicobacter pylori.

XX PN W09737044-A1.

XX PD 09-OCT-1997.

XX PF 27-MAR-1997; 97WO-US05223.

XX PR 06-DEC-1996; 96US-0761318.

XX PR 29-MAR-1996; 96US-0625811.

XX PR 02-APR-1996; 96US-0758731.

XX PR 25-OCT-1996; 96US-0736905.

XX PR 28-OCT-1996; 96US-0738859.

XX PA (ASTR) ASTRA AB.

XX PA Alm RA, Smith D;

XX DR WPI; 1997-503122/46.

XX DR N-PSDB; AAW24985.

XX PT Helicobacter pylori nucleic acid sequences and encoded

PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori

PT infection and for diagnosis of H. pylori infection

XX Claim 14; Pages 780-781; 1145pp; English.

XX This sequence is a H. pylori protein of unspecified function.

CC The protein may be used in a vaccine to prevent or treat H. pylori

CC infection or to identify H. pylori polypeptide binding compounds,

CC useful as potential H. pylori life cycle activators or inhibitors. The

CC DNA and probes derived from it may be used for the identification of

CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic

CC acid sequences complementary to the DNA act as antisense sequences and

CC can be used to prevent the translation of H. pylori mRNA. Antibodies

CC against the protein can be used in immunoassays to evaluate the abundance

CC and distribution of H. pylori-specific antigens. The genomic sequence of

CC H. pylori (ATCC 35679) was determined from overlapping contigs generated

CC by mechanically shearing the bacterial DNA. The sequences were analysed

CC for ORF of at least 180 nucleotides, and the predicted coding regions

CC defined by computer evaluation. To identify likely H. pylori antigens for

CC vaccine development, the amino acid sequences predicted from various ORF

CC were analysed for significant homology to other known or exported

CC membrane proteins. Having identified and determined the sequences of

CC interest, particular regions can be isolated from H. pylori by PCR

CC amplification for recombinant polypeptide production, e.g. in E. coli

CC hosts.

XX Sequence 660 AA;

Qy Query Match 17.9%; Score 625; DB 18; Length 660;

Best Local Similarity 28.2%; Pred. No. 4.3e-35;

Matches 185; Conservative 129; Mismatches 255; Indels 86; Gaps 22;

Qy 1 KIYDNKNOLIADL-GSERRNAQANDIPTDLVKAIVSIEDHRRFDRGIDTRIILGAFELR 59

Db 48 qildrkrlianiydkerfyarfeepfvesllavedltlffhgghindavnamrak 107

Qy 60 NLOSNS-LOGGSTLTQOLIKLTYFSTSDOTISKQAEAWLAIOLEKATKQEILTYYI 118

Db 108 naksgryteggstltqqlvkmvl---trektitrlkealisiriekvlseilleryi 164

Qy 119 NKVYMSNGYMGQTAQNYGKDLNLSLPQLALLAGPQAPNOYDPYSHPEAADRRNL 178

Db 165 nqtffghygyvktaslgfyfkkpdkldlklkeitmvalprapsfydptknlefsrand 224

Qy 179 VLSEMKNGYISAEQYKAVN-TPITDGLQSLKASNPAYMDNLYKEVINQVEEETGYN 237

Db 225 llrrlyslgwissnelksalnepivynqtstgniapy-----vdevlklqldlg-- 276

Qy 238 LLTTCMDVYTNVDQEAQKHLWDIY-----NTDEYVAYPDDELQVASTIV 281

Db 277 lktggytkltdldyqrlaleslrfghqkilekiaekpktnasndkdednlnasmivt 336

Qy 282 DVSNGKVIQALGARHSSNVSGFNGQAVETNRDWSGTMKPTTDYAPALEYGVYDSTATIV 341

Db 337 ectstgkialalvggidykks---afnratqakrfgsaakpfv-yqlafdnngy--sttski 390

Qy 342 HDEPNYPTGTPVY-----NMDRGYFGNITLQYALQOQSRNPVAVETLNKVG 388

Db 391 pdtarnfengn---ysknsqvphawhpsnytrkfglvtlqlealslnlatinsldqgl 447

Qy 399 LNRKATFUNGIGIDYPSIHYSHNAISSNTTESDKKYGASSEKMAAAYAAFANGGYTKPMY 448

Db 448 fekiyqslsdmgfknlpkdlslvigs-----faispidaaeakyslfsnygtmklpml 499

Qy 449 IHKVVFSDSSEKFSNVGTRAM-KETAYMMTDMKTKVLTGTGRNAYLAWLPOAGKTGT 507

Db 500 iesitnqqnevktftptetkkitkeqafitlsalmdavengtgsarikgleiagktgt 559

Qy 508 SNYTDDEEIEENHIKTSQFVAPDELFGYTRKYSMAVWTGYSNRLTPLVGNGLT---VAARKV 564

Db 560 sn-----nni-----dawfigftptlqslvifgrddn--tp-igkgatggvvsapv 602

Qy 565 YRSMNTYL--SEGSNPEDWNIPEGLYNGEFVFKNGARSTWNSPAPQPPSTESS 617

Db 603 ysyfmrllalepslkrkfdvpglirk--eivdkipyyspnsitp-tpkktdds 654

Search completed: June 13, 2002, 08:42:39
Job time: 125 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2002, 08:41:44 ; Search time 13.62 seconds
(without alignments)
1893.335 Million cell updates/sec

Title: US-08-961-083-2

Perfect score: 3484

Sequence: 1 KIYDNKNLIADLGSRVYN.....TOGSNTPTDQONPOPAQP 666

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3484	100.0	719	1	BPBA_STRPN
2	2698.5	77.5	637	1	BPBA_STROR
3	999	28.7	914	1	BPBA_BACSU
4	739	21.2	714	1	BPBF_BACSU
5	712	20.4	726	1	BPBA_AQUAE
6	680.5	19.5	798	1	BPBA_NEIFL
7	642.5	18.4	822	1	BPBA_PSEAE
8	640	18.4	798	1	BPBA_NEIFA
9	640	18.4	798	1	BPBA_NEIMA
10	635.5	18.2	797	1	BPBA_NEIGO
11	633	18.2	798	1	BPBA_NEICI
12	617	17.7	825	1	BPBA_VIBCH
13	598.5	17.2	624	1	BPBD_BACSU
14	591.5	17.0	853	1	BPBA_HAEIN
15	584	16.8	777	1	BPBP_VIBCH
16	567	16.3	850	1	BPBA_ECOLI
17	526	15.1	844	1	BPBP_ECOLI
18	525.5	15.1	781	1	BPBP_HAEIN
19	502.5	14.4	787	1	BPBA_RICPR
20	490.5	14.1	809	1	BPBA_XYLFA
21	472	13.5	760	1	BPBP_BUCAI
22	450.5	12.9	770	1	BPBC_ECOLI
23	312.5	9.0	207	1	TRG_ALCEU
24	207	5.9	716	1	BPBP_BACSU
25	206	5.9	233	1	MTGA_NEIGO
26	206	5.9	645	1	SP5D_BACSU
27	194.5	5.6	224	1	MTGA_ACICA
28	193.5	5.6	598	1	FTSI_MESVI
29	192	5.5	233	1	MTGA_NEIMA
30	185.5	5.3	242	1	MTGA_KLEPN
31	183.5	5.3	242	1	MTGA_ECOLI
32	181	5.2	246	1	MTGA_HAEIN
33	164	4.7	588	1	FTSI_ECOLI

34	147	4.2	1861	1	APU_THETU	P38536 t amylopull
35	142.5	4.1	750	1	BPBX_STRPN	P14677 streptococc
36	142	4.1	610	1	FTSI_HAEIN	P45059 haemophilus
37	141	4.0	716	1	YQGF_BACSU	P54488 bacillus su
38	141	4.0	2843	1	APC_HUMAN	P25054 homo sapien
39	140.5	4.0	516	1	P34_ENTFC	P13692 enterococcu
40	139.5	4.0	706	1	PLB2_YEAST	P33653 streptococc
41	138	4.0	1337	1	DEXT_STRDO	P39653 saccharomyc
42	136.5	3.9	567	1	CH13_CANAL	P40954 candida alb
43	136	3.9	579	1	FTSI_BUCAI	P57317 buchnera ap
44	135.5	3.9	1007	1	Y741_CHLMU	O9pj16 chlamydia m
45	135	3.9	1365	1	GTFS_STRDO	P29336 streptococc

ALIGNMENTS

RESULT 1					
ID	BPBA_STRPN	STANDARD;	PRT;	719 AA.	
AC	Q04707;				
DT	01-OCT-1994 (Rel. 30, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Penicillin-binding protein 1A (BPB-1A) (Exported protein 2).				
GN	PONA OR EXP2 OR SP0369.				
OS	Streptococcus pneumoniae.				
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1313;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=45607, AND 63915;				
RX	MEDLINE=93010977; PubMed=1396576;				
RA	Martin C., Sibold C., Hakenbeck R.;				
RT	"Relatedness of penicillin-binding protein 1a genes from different clones of penicillin-resistant Streptococcus pneumoniae isolated in South Africa and Spain."				
RL	EMBO J. 11:3831-3836(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=R6;				
RX	MEDLINE=92325042; PubMed=1624444;				
RA	Martin C., Briese T., Hakenbeck R.;				
RT	"Nucleotide sequences of genes encoding penicillin-binding proteins from Streptococcus pneumoniae and Streptococcus oralis with high homology to Escherichia coli penicillin-binding proteins 1a and 1b."				
RL	J. Bacteriol. 174:4517-4523(1992).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=TIGR4;				
RX	MEDLINE=21357209; PubMed=11463916;				
RA	Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D., Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J., Durkin A.S., White M., Kolonay J.F., Nelson W.C., Peterson J.D., Umayam L.A., Whitt O., Salzberg S.L., Lewis M.R., Radune D., Holtzaple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L., McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;				
RT	"Complete genome sequence of a virulent isolate of Streptococcus pneumoniae."				
RL	Science 293:498-506(2001).				
RN	[4]				
RP	SEQUENCE OF 293-369 FROM N.A.				
RC	STRAIN=R6X;				
RX	MEDLINE=95020625; PubMed=7934910;				
RA	Pearce B.J., Yin Y.B., Masure H.R.;				
RT	"Genetic identification of exported proteins in Streptococcus pneumoniae."				
RL	Mol. Microbiol. 9:1037-1050(1993).				
CC	!- FUNCTION: CELL WALL FORMATION.				
CC	!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.				

CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: THESE STRAINS ARE PENICILLIN-SENSITIVE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67873; CAA48073.1; -
DR EMBL: X67872; CAA48072.1; -
DR EMBL: M90527; AAA26956.1; -
DR EMBL: AE007349; AAK74536.1; -
DR TIGR: S28038; S28038.
DR TIGR: SP0369; -
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme; Complete proteome.
FT ACT_SITE 370 370 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT DOMAIN 658 683 SER-RICH.
FT VARIANT 124 124 T -> A (IN STRAIN R6).
FT VARIANT 386 386 V -> I (IN STRAIN 63915).
FT VARIANT 388 388 D -> E (IN STRAIN R6).
FT VARIANT 397 397 E -> K (IN STRAIN 63915).
FT VARIANT 523 523 M -> I (IN STRAIN 63915).
FT VARIANT 533 533 D -> E (IN STRAIN 45607).
FT VARIANT 540 540 T -> S (IN STRAINS 63915 AND R6).
FT VARIANT 657 657 N -> S (IN STRAINS 45607 AND R6).
SQ SEQUENCE 719 AA; 79758 MW; 59D397E83B4B3AA6 CRC64;

Query Match 100.0%; Score 3484; DB 1; Length 719;
Best Local Similarity 100.0%; Pred. No. 2.5e-193;
Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KIYDNKQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFPHRGDTIRILGAFNRN 60
DB 54 KIYDNKQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFPHRGDTIRILGAFNRN 113
QY 61 LQSNLSGGSTLTQOLIKLTFTSTDSQITSRKAQEAWLAIQLEQKATKQEIITYINK 120
DB 114 LQSNLSGGSTLTQOLIKLTFTSTDSQITSRKAQEAWLAIQLEQKATKQEIITYINK 173
QY 121 VYMSNGYGMQTAQNYGKDLNLSLPQLALLAGMPQAPNQYDPYSHPEAAQDRRLVL 180
DB 174 VYMSNGYGMQTAQNYGKDLNLSLPQLALLAGMPQAPNQYDPYSHPEAAQDRRLVL 233
QY 181 SEMNKGVIISAEQYKAVNTPTDGLQSLKSASNPAYMDNLYKEVINQVEEETCYNLLT 240
DB 234 SEMNKGVIISAEQYKAVNTPTDGLQSLKSASNPAYMDNLYKEVINQVEEETCYNLLT 293
QY 241 TGMVYTNVDQEAQKHLWDIYNTDEYVAYPDELOQVASTIVDVSNGKVIQALGARHOSN 300
DB 294 TGMVYTNVDQEAQKHLWDIYNTDEYVAYPDELOQVASTIVDVSNGKVIQALGARHOSN 353
QY 301 VSFQINQAVETNRDWSGTMKPTIDYAPALEYGVYDSTATIVHDEPYNYPGNTPVNMDR 360
DB 354 VSFQINQAVETNRDWSGTMKPTIDYAPALEYGVYDSTATIVHDEPYNYPGNTPVNMDR 413
QY 361 GYFGNITLOYALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHSNAISSNTTESD 420
DB 414 GYFGNITLOYALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHSNAISSNTTESD 473
QY 421 KKYGASSEKMAAAYAAAFANGGTYTKPMYIHKVWFSDGSEKESFNVGTFRAMKETTAYMMTD 480
DB 474 KKYGASSEKMAAAYAAAFANGGTYTKPMYIHKVWFSDGSEKESFNVGTFRAMKETTAYMMTD 533

QY 481 MMKTVLTGTCRNAYLAWLPQAGKTGTSNYTDEEIEHNHIKTSQFVAPDELFAGYTRKYSM 540
DB 534 MMKTVLTGTCRNAYLAWLPQAGKTGTSNYTDEEIEHNHIKTSQFVAPDELFAGYTRKYSM 593
QY 541 AVWTGYSNRLTPLVGNGLTVAAKVYRSMTYLSGSGNPEDWNIPEGLYRNGEEYFKNGAR 600
DB 594 AVWTGYSNRLTPLVGNGLTVAAKVYRSMTYLSGSGNPEDWNIPEGLYRNGEEYFKNGAR 653
QY 601 STWNSPAPQPPSTESSSSSSSSSTSSSTSTPSTNNSTTTNNNTQOOSNTTTPDQOQNON 660
DB 654 STWNSPAPQPPSTESSSSSSSSSTSSSTSTPSTNNSTTTNNNTQOOSNTTTPDQOQNON 713
QY 661 POPAQ 666
DB 714 POPAQ 719

RESULT 2

BPAPA_STOR ID BPAPA_STOR STANDARD; PRT; 637 AA.
AC Q00573;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1A) (Fragment).
GN PONA.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92325042; PubMed=1624444;
RA Martin C., Briese T., Hakenbeck R.;
RT "Nucleotide sequences of genes encoding penicillin-binding proteins
RT from Streptococcus pneumoniae and Streptococcus oralis with high
RT homology to Escherichia coli penicillin-binding proteins 1a and 1b.";
RL J. Bacteriol. 174:4517-4523(1992).
CC -!- FUNCTION: CELL WALL FORMATION.
CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M90528; AAA26958.1; -
DR PIR: B42893; B42893.
DR InterPro: IPR001264; Transglycosyl.
DR InterPro: IPR001460; Transpeptidase.
DR Pfam: PF00912; Transglycosyl; 1.
DR Pfam: PF00905; Transpeptidase; 1.
DR ProDom: PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Antibiotic resistance; Cell wall;
KW Multifunctional enzyme.
FT ACT_SITE 371 371 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT NON_TER 637 637
SQ SEQUENCE 637 AA; 70891 MW; A6D198BCEA603A63 CRC64;

Query Match 77.5%; Score 2698.5; DB 1; Length 637;
Best Local Similarity 87.5%; Pred. No. 3.5e-148;
Matches 511; Conservative 39; Mismatches 33; Indels 1; Gaps 1;

QY 1 KIYDNKQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFPHRGDTIRILGAFNRN 60
DB 54 KIYDNKQIADLGSERRVNAQANDIPTDLVKAIVSIEDHRRFPHRGDTIRILGAFNRN 113
QY 61 LQ-SNSLOGGSTLTQOLIKLTFTSTDSQITSRKAQEAWLAIQLEQKATKQEIITYINK 119

```
Db 114 LRGGGGLOCASTLTQOLIKLYFTSTSDQTLRSRAQEAWLAVOLQKATQKEILTYVIN 173
Qy 120 KYVNSNGYGMOTAAQNYVYGDNLNLSIPQALLAGMPOAPNOYDPYSHPEAAQDRNLV 179
Db 174 KYVNSNGYGMOTAAQNYVYGDNLNLSIPQALLAGMPOAPNOYDPYSHPEAAQDRNLV 233
Qy 180 LSEKNQGYISAEQYEKAVNTPTDGLQSLKASASNPAYMDNLYKEVINQVEEETGYNLL 239
Db 234 LSEMKGGGYITAEQYEKAVNTPTDGLQSLKASASNPAYMDNLYKEVINQVEEETGYNLL 293
Qy 240 TTGMVYNNVDQEAQKHLWDIYNTDEYVAYPDDELQVASTVDVNSGVKVIQAOLGARHQS 299
Db 294 TTGMVYNNVDQEAQKHLWDIYNTDEYVAYPDDELQVASTVDVNSGVKVIQAOLGARHQS 353
Qy 300 NVSFGINGAETNRDWSGTMKPTDYAPALEYGVYDSTATVHDSPYNYPGTNTPTVYND 359
Db 354 NVSFGINGAETNRDWSGTMKPTDYAPALEYGVYDSTATVHDSPYNYPGTNTPTVYND 413
Qy 360 RGYFGNITLQYALQOSNRVAVETLNKVLNRAKTFNLGLGIDYPSIHYNAISSNTTSES 419
Db 414 KSYFGNITLQYALQOSNRVAVETLNKVLNRAKTFNLGLGIDYPSIHYNAISSNTTSES 473
Qy 420 DKYGGASSEKMAAAYAAPANGTYTKPMYIKHVPSDGESEKFEFNVGTRAMKETAYMMT 479
Db 474 DKYGGASSEKMAAAYAAPANGTYTKPMYIKHVPSDGESEKFEFNVGTRAMKETAYMMT 533
Qy 480 DMKKTLYTGYGRNAYLAWLPQAGKTGTSNVTDEIENHIKTSOFVAPDELPAQVTRKYS 539
Db 534 DMKKTLYTGYGRNAYLAWLPQAGKTGTSNVTDEIENHIKTSOFVAPDELPAQVTRKYS 593
Qy 540 MAVTGYNSRLTPLVNGGLTVAAKYRSMYTLSEGSNPEDWNI 583
Db 594 MAVTGYNSRLTPIVGDGYFAAKYRSMYTLSEDDHPGDWTM 637

RESULT 3
PBPA_BACSU STANDARD; PRT; 914 AA.
AC P39793;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1A/1B (PBP1) [Includes: Penicillin-
DE insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase);
DE penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].
GN PONA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 499-515.
RC STRAIN=168;
RX MEDLINE=95113769; PubMed=7814321;
RA Popham D.L., Setlow P.;
RT "Cloning, nucleotide sequence, and mutagenesis of the Bacillus
RT subtilis ponA operon, which codes for penicillin-binding protein
RT (PBP) 1 and a PBP-related factor.";
RL J. Bacteriol. 177:326-335(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / MARRBURG;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serrif P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RT the serA and kds loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
RN [3]
RP GROWTH REQUIREMENTS.
RC STRAIN=168;
RX MEDLINE=98389671; PubMed=9721295;
```

```
RA Murray T., Popham D.L., Setlow P.;
RT "Bacillus subtilis cells lacking penicillin-binding protein 1 require
RT increased levels of divalent cations for growth.";
RL J. Bacteriol. 180:4555-4563(1998).
RN [4]
RP SUBCELLULAR LOCATION.
RC STRAIN=168;
RX MEDLINE=99255546; PubMed=10322023;
RA Pedersen L.B., Angert E.R., Setlow P.;
RT "Septal localization of penicillin-binding protein 1 in Bacillus
RT subtilis.";
RL J. Bacteriol. 181:3201-3211(1999).
CC -|- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNITS) (BY SIMILARITY).
CC -|- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -|- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN, PROBABLY FOUND ALL
CC OVER THE WHOLE CELL AT LOW CONCENTRATIONS. ALSO LOCALIZES TO THE
CC DIVISION SITE IN VEGETATIVE CELLS.
CC -|- DEVELOPMENTAL STAGE: EXPRESSION IS CONSTANT DURING GROWTH,
CC DECREASES DURING SPORULATION AND IS INDUCED APPROXIMATELY 15 MIN
CC INTO SPORE GERMINATION.
CC -|- PTM: THE PRODUCT EXPRESSED FROM THE TRANSLATION OF THE PONA GENE
CC APPEARS AS TWO BANDS ON A GEL (1A AND 1b), BUT THE SPECIFIC AMINO
CC ACID SEQUENCE OF EACH PROTEIN IS UNKNOWN.
CC -|- MISCELLANEOUS: CELLS LACKING THE PROTEIN REQUIRE INCREASED LEVELS
CC OF MG2+ OR CA2+ FOR GROWTH AND GERMINATION. APPROXIMATELY 50% OF
CC CELLS WITHOUT THE PROTEIN CONTAIN ABNORMAL FTSZ RINGS, SUGGESTING
CC IT IS INVOLVED IN SEPTUM SYNTHESIS; INCREASED LEVELS OF MG2+ OR
CC CA2+ ONLY PARTIALLY ELIMINATE THE SEPTATION DEFECTS.
CC -|- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -|- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U11883; AAA64947.1; -.
CC EMBL; L47838; AAB38459.1; -.
CC EMBL; Z99115; CAB14148.1; -.
CC HSSP; P02751; 1FNA.
CC Subtilist; BG10954; ponA.
CC InterPro; IPR003961; FN.III.
CC InterPro; IPR001264; Transglycosyl.
CC InterPro; IPR001460; Transpeptidse.
CC Pfam; PF00041; fn3; 1.
CC Pfam; PF00912; Transglycosyl; 1.
CC Pfam; PF00905; Transpeptidase; 1.
CC ProDom; PD001895; Transglycosyl; 1.
CC SMART; SM00060; FN3; 1.
CC Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
CC Hydrolase; Multifunctional enzyme; Transmembrane; Signal-anchor;
CC Antibiotic resistance; Complete proteome.
CC DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 38 58 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
CC POTENTIAL.
CC DOMAIN 59 914 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 77 246 TRANSGLYCOSYLASE.
CC DOMAIN 329 662 TRANSEPTIDASE.
CC FT ACT_SITE 390 390 ACYLATED BY PENICILLIN (BY SIMILARITY).
CC SQ SEQUENCE 914 AA; 99562 MW; 6978E33DFE2423B6 CRC64;
```

Query Match 28.7% Score 999; DB 1; Length 914;
 Best Local Similarity 31.2%; Pred. No. 4.2e-50;
 Matches 254; Conservative 131; Mismatches 256; Indels 174; Gaps 24;

Qy 2 IYDNKNQIADLGSERRVNAQANDIPTDLVKAIVSIEDHFRFDHGTIDTILGAFNL 61
 Db 79 IYDNKNGEIAVGAERKRYVSEIDPDVVKVKAFTATEDARFYEHHGIDPVVRIGALVANF 138
 Qy 62 QSN-SLOGGSLTLOOLIKLTFTSTSDQTSRKAQEAWLAIQLEOKATQBEILTYINK 120
 Db 139 KDGCAEGSGTITQOVVANSLLS---HOKTLKRKVOEWLSIQLEARNYSKDEILEMYLNR 195
 Qy 121 VYMSNGNYGMOAQNYYG-KDLNLSLPQALLAGMPQAPNQYDPYSHPEAAQDRRLV 179
 Db 196 IYFSPRAYIGKAAEEFPGVTDLSKLTVEQAATLAGMPQSTAYNPVKNPDKAEKRRNV 255
 Qy 180 LSEKNQGYISAEQYKAVNPITD-GLQSLK-----SASNPAYMDNLYKEVINQVEE 233
 Db 256 LSLMKKQGFISDSQYNKAKKAVKDEGVYSQYKESASTNYSAPVE---EVMKEIDEK 311
 Qy 234 TGYNLLTGMVDYTNVDOEAOKHLWDIYNTDEYVAYPDDELQVASTIVDSNGKVIQALG 293
 Db 312 SDVPSADGLKITYTLPDKAOKLDELMDGOT-VGFTEG-MQGGVTLTDTKNGEVRALGA 369
 Qy 294 ARHOSNYSFGINQAVETNRDWSMTKPIDYAPALEYGVYDSTATIVHDPEYNPYPTNT 353
 Db 370 GRNOPVG---GFNATQTKAOPGSIKPIIDYGPVIEKNK-STVEQIDDSAYTY-SNGK 424
 Qy 354 PYNWDRGYFNGNITLOALQOSRNPVAVETLNKVLNKLAKTFLNGLGIDYPSIHSNAIS 413
 Db 425 PIRMDRKYLPISMRVYALASRNIPALUKFAQVAGKDTAVDFANGLGLTKDNVTEAYS 484
 Qy 414 SNTTESDKYKASSSEKMAAAFAANGTYKPKMYIKHWFSDGSEKFSNVGTRAMKET 473
 Db 485 IGGFGND--GVSPITMAGVSAFCNNGTYNEPHFKVSEIENGDGTLDTPKSKSAMS 542
 Qy 474 TAYMTDMKTVLYGTGRNAYLAWLPOAGKTGTSNYTDEIEHNIKTSQVAPDELPAFAG 533
 Db 543 TAFMITDLKTAVKTGQLAQVQVEVAGKTGTTNFDNEVKRY-NIASGGARDSWFVG 601
 Qy 534 YTRKYSMAVMTGYSNR---LPLVNGLTVAKVYRSMNTYLSGSGS---NPEDWNIEPG 586
 Db 602 YTPQITAAVMTGMGENEAGKSLSAEOKVAKRIFAQLIADVDGSGSFERPD--SVVEA 659
 Qy 587 LYRNG-----BEVFNKGARST----- 602
 Db 660 TVEKSNPAKLAPNTPSKDKLTVEFYKGTAPSTVSKTYEKEEETAKLSGLNVKYDKD 719
 Qy 603 -----W-----NSPAPQ-----QPPS----- 613
 Db 720 NQSLTSLWNYDGDATFAVKQSVGSGSYSEIQNSSAKEAVISGVOPGYSVKYKFEVTA 779
 Qy 614 -----TESSSSSSSSSTSQSSSTTPTST 635
 Db 780 KSTASTSYEVKPAEDEDKQKQOQDDEKQDEKTDQDDTQDDSQKDGQDQDDTDDST 839
 Qy 636 N-----NSTTTNP--NNNTQOSNTTPDQONQNPQ 662
 Db 840 NDQDKKQDNTNTFSDNNNQDSNDN-DNDNSNQ 873

RESULT 4

BPPE_BACSU ID BPPE_BACSU STANDARD; PRT; 714 AA.
 AC P38050;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Penicillin-binding protein 1F (PBP-1F).
 GN PBP OR PONA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.

NCBI_TaxID=1423;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98240224; PubMed=9579061;
 RA Noback M.A., Holsappel S., Kiewiet R., Terpstra P., Wambutt R.,
 RA Wedler H., Venema G., Bron S.;
 RT "The 172 kb prkA-addAB region from 83 degrees to 97 degrees of the
 RT Bacillus subtilis chromosome contains several dysfunctional genes,
 RT the glyB marker, many genes encoding transporter proteins, and the
 RT ubiquitous hit gene.";
 RL Microbiology 144:859-875(1998).
 [2]
 RP SEQUENCE OF 1-129 FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=93328693; PubMed=8335642;
 RA Popham D.L., Setlow P.;
 RT "Cloning, nucleotide sequence, and regulation of the Bacillus
 RT subtilis pbpF gene, which codes for a putative class A
 RT high-molecular-weight penicillin-binding protein.";
 RL J. Bacteriol. 175:4870-4876(1993).
 [3]
 RP SEQUENCE OF 122-714 FROM N.A.
 RX MEDLINE=9304140; PubMed=1459957;
 RA Hansson M., Hedersstedt L.;
 RT "Cloning and characterization of the Bacillus subtilis hemery gene
 RT cluster, which encodes protoheme IX biosynthetic enzymes.";
 RL J. Bacteriol. 174:8081-8093(1992).
 CC -!- FUNCTION: CELL WALL FORMATION. MAY BE INVOLVED IN OUTGROWTH OF THE
 CC GERMINATED SPORE OR IT COULD FUNCTION IN THE SYNTHESIS OF THE GERM
 CC CELL WALL.
 CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Membrane-associated.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSION REMAINS CONSTANT DURING VEGETATIVE
 CC GROWTH, DECREASES DURING EARLY SPOULATION, AND IS INDUCED IN THE
 CC FORESPORE DURING LATE SPOULATION.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSEPTIDASE FAMILY.
 CC PROTEINS
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Y14083; CAA74517.1; -;
 CC EMBL; L10630; AAA71942.1; -;
 CC EMBL; M97208; AAA22516.1; -;
 CC EMBL; Z99109; CAB12851.1; -;
 CC PIR; A40614; A40614.
 CC Subtilisin; BG10428; pbpF.
 CC InterPro; IPR001264; Transglycosyl.
 CC InterPro; IPR001460; Transpeptidase.
 CC Pfam; PF00912; Transglycosyl; 1.
 CC Pfam; PF00905; Transpeptidase; 1.
 CC ProDom; PD001895; transglycosyl; 1.
 CC Peptidoglycan synthesis; Cell wall; Transmembrane; Signal-anchor;
 CC Complete proteome.
 FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 13 33 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
 FT POTENTIAL.
 FT DOMAIN 34 714 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 49 217 TRANSGLYCOSYLASE.
 FT DOMAIN 297 592 TRANSEPTIDASE.
 FT ACT SITE 359 ACYLATED BY PENICILLIN (BY SIMILARITY).
 FT SEQUENCE 714 AA; 79278 MW; 08D96718C84BB434 CRC64;

Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
[2]

[illegible]

Qy	61	LQSNLSQ--GGSTLTQQLIKLTSTSTSDQTSRKAQEAHLAIQLQEKATKQEIITYYIN	119
Db	109	LKSHIQTGGSTTMOQAK--NYF--LTNRSFRRKLTNEILLALQIERQTLKDEILELYN	165
Qy	120	KVYMSNGYMGOTAAQNYGCKDLNNLSLPOLALLACGMPQAPNOYPDYPSHPEAAODRRNLV	179
Db	166	KIYLGNRAYGIEAAQVYIGKPLDUSLAEAMIAGLPKAPSRYNPLVNPTSTERRNWI	225
Qy	180	LSEMKNGYISAEQYKAVNTPITDGLQLSLKASNPAYMDNYLKVINQVBEET--G	236
Db	226	LERMLKLGFIQQRYQAQVEEPINASVHVQTPELNPAPYIAEMARAEWVGRYGEAYTEG	285
Qy	237	NLLTGMDEVTN-----VDQEAQ-----KHL-----	257
Db	286	KVITTVRSDLQNAASQSVRDGLDIYDQRHGIRGYPETRLPGQTRDAWLKLGQORSGIGLE	345
Qy	258	-----WD-----IYNTDEVVAYPDEQLQV	276
Db	346	PAIVTQVEKSGIWMTRDGEAAVTWDSMKWARPFLSNNSMGPMRQPADVAQAQGLQV	405
Qy	277	-----ASTIVDVSGKVIQAQLGA--RHQSSNVSFGINQAVETNRDWS	317
Db	406	QROEDGTLRFVQIPAAQSALISLDPKDGAIRSLVGGSFSEQSN---YNRAIQAKRQPCS	461
Qy	318	TMKPIDIYAPALEYGVYDSTATIVDHEPYN-----PGTNPVYNWDRGYGN	365
Db	462	SFKPFI--YSAALONGF--TAASLVNDAPIVFVDEYLDKVMRPKNDTNT-----FLGP	510
Qy	366	ITLQYALQOGRNPVAVETLNKVLGNRAKFLNGLGIDYPSI--HYSNAISSNTTESDKKY	423
Db	511	IPLREALYKGRNMVSIIRVLOGLGIERAISVITKFGFQDELPRNFSALGLGTAT-----	563
Qy	424	GASSEKMAAIAAFANGFTYKPMWTHKVVFSQ-----SEKEFSN	464
Db	564	-VTMETAGAWSVPANGCYKVNYPVIERIESDQVLYQANPPRVPEQOVAADAEDAGN	622
Qy	465	VG-----TRAMKETTAYMMTMMKTLVTYGTGR	492
Db	623	PGDPEHPESAEGEGSIEAQOVAAKAQTTEPTPEARIIDARTAYIMTSLQDVIKRGTR	682
Qy	493	NAY--LAWLPQANGTGTSNYTDDEEIHKTSSQFVAPDELPAFYTRYKYSMAVWTGYSNRLT	551
Db	683	RALAKRKTDLAGTGTGTN-----DSKDGWFSGYNSDYVTSVMWGFQDQPET	727
Qy	552	PLVGN--GLTVAAKVYRSMWYLSGCSNPEDWNIPGELYRNGEFVKNGARSTWNSPAP	608
Db	728	--LGRREYGGTVALPIWIRYMGALCKDKPMHTMAEPPGI-----VSLRIDPVTGRSAAP	779
Qy	609	QQPSTESSSSSDSTSSQSSSTTPSTN	636
Db	780	GTPGAYFEMFKNE-----TPPSVN	799
RESULT 8			
ID	BPBA_NEILA	STANDARD;	PRT; 798 AA.
AC	087579;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DE	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Penicillin-binding protein 1A (PBP-1a) (BPPIa) [Includes: Penicillin-		
DE	insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase);		
DE	penicillin-sensitive transpeptidase (EC 3.4.-) (DD-transpeptidase)]		
GN	MCA OR PONA.		
OS	Neisseria lactamica.		
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
OX	NCBI_TaxID=486;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NRL 3716;		
RA	Ropp P.A., Nicholas R.A.;		
RT	"Nucleotide sequence of the ponA gene encoding penicillin-binding		
RT	protein 1 from Neisseria lactamica."		

Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
 CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
 CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
 CC SUBUNITS) (BY SIMILARITY).
 CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
 CC similarity).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSEPTIDASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF085689; AAC35363.1; -
 CC InterPro; IPR001264; Transglycosyl.
 CC InterPro; IPR001460; Transpeptidase.
 CC Pfam; PF00912; Transglycosyl; 1.
 CC Pfam; PF00905; Transpeptidase; 1.
 CC ProDom; PD001895; Transglycosyl; 1.
 CC Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
 CC Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
 CC Signal-anchor; Antibiotic resistance.
 CC DOMAIN 1 9
 CC TRANSMEM 10 30
 CC SIGNAL-ANCHOR (POTENTIAL).
 CC (POTENTIAL).
 CC PERIPLASMIC (POTENTIAL).
 CC TRANSGLYCOSYLASE.
 CC DOMAIN 31 798
 CC DOMAIN 50 218
 CC TRANSEPTIDASE.
 CC DOMAIN 378 700
 CC ACT SITE 461 461
 CC ACYLATED BY PENICILLIN (BY SIMILARITY).
 CC SEQUENCE 798 AA; 88108 MW; 3B9C7672886935D6 CRC64;
 SQ
 Query Match 18.4%; Score 640; DB 1; Length 798;
 Best Local Similarity 26.0%; Pred. NO. 1.6e-29;
 Matches 203; Conservative 126; Mismatches 265; Indels 186; Gaps 24;
 QY 2 IYDNKNQLIADLGSRRVNAQNDIPTDLVKAIVSIEDHRRFDHGDITIRILGAFNL 61
 DB 52 VYSADGKIIGYGEORREFTKIGDFPEVLRNAVIAEDKREYQHWGVWGVAVVNGNI 111
 QY 62 QSNLQ-QGSTLTQOLIKLTVFSTSDQTSRKAQEAWLAIQLEKQKATQOEILTYINK 120
 DB 1.12 VSGSQSGASTITQVANKFYL-----SEKTRKNEALLAYKIEQSLSKKILELYNQ 168
 QY 121 VYMSNGYGMOTAAQNYGKDLNNLPLQALLAGMPQAPNOYDPSHPAEAQRRLNLV 180
 DB 169 IYLGORAYGEASAAQIYFNKQVDRLTAEAAMLAGLRAPSAYNPIVNERAKLRQKYL 228
 QY 181 SEMKNQGIISABOYEKAVNTPTDGLQSLKSNAPYADWN-----YLKEVINO-VVEE 233
 DB 229 NNMLEKMITLQORDQALNEEL-----HYERFVQKIDQSLAYVAMVRQELYEK 277
 QY 234 TGYNLTTGMDVYTNV-----DQEAQ--- 254
 DB 278 YGEDAYTOGLKVYTVYTRTDHQKAATEALKARNFDRGSSYGAESYIDLGRDEDAEAV 337
 QY 255 -KHLWDIYNTDEYV-----AYPDDE----- 273
 DB 338 SQYLSGLTVDMKWPVAVLDVTKKNNVVITQLPGGKRVTLDRRALGFAARAVDNKMGEDR 397
 QY 274 -----IQVASTIVDSNGKVIQALGARHQSSNVSFGINQAVE 310
 DB 398 IRRGAVIRVRNNGRWAVVQEPPLQALVSLDAKTGAVRALVGVGYDFHSKT---FNRAIQ 454

QY 311 TNRMGSTMKPTIDYAPALEYGVYDSTATIVHDEPNYPGNT-----TPVYNWDRGVFG 364
 DB 455 AMRQPGSTFKPEV-YSAALSQGM--TASTINDAPISLPGKPGNSVWTP-KNSDGRYSG 510
 QY 365 NITLOALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSHYSHYNAISSNTTESDKKVG 424
 DB 511 YITLQALTSKNMYSIRILMSIGVYAAQYIRRRFGKPSLPSLSMALGTGET----- 565
 QY 425 ASSEKMAAAYAAFANGGTYIKVPMYTHKVVFSQSGSEK-----FSNVGTRAMKETAYMM 478
 DB 566 -TPLRIAEAYSVFANGGRVRSYVTDKIYDSEGRRAQMQPLVAGONAPQADIPRDAYIM 624
 QY 479 TDMKTKVLTGYGTGRNA-YLAWLPQAGKTGTSNYTDEETENHIKTSQFVAPDEFALFYRK 537
 DB 625 YKIMQDVVRVGTARGASALGRSDIAGKTGTTN-----DN-----KDAWFGVFNPD 669
 QY 538 YSMVAVWTGYSN-RLTPLVNGNLTVAAKVYRSMTYLSGSGNPEDWNIPBGLYR-NGEYVF 595
 DB 670 VVTAVYIGFDKPKSMGRAGYGTIAVPVWVDYMRFAKLRPGKMKMPDGVVAGNGEYTM 729
 QY 596 KN-----GARSTWNSPAPQPPS-----TESSSSSDSSTSSSTSTPSTNNST 639
 DB 730 KEHMTVDPLGLMDNGGAAPQ---PSRRVKEDDGGAAEGGRQEADESRQDMQETPLVPSNT 787
 RESULT 9
 ID BPBA_NEIMA STANDARD; PRT; 798 AA.
 AC 005194;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Penicillin-binding protein 1A (PBP-1a) (PBP1a) [Includes: Penicillin-
 DE insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase);
 DE Penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].
 DE MRCA OR PONA OR NMA0655 OR NMB1807.
 OS Neisseria meningitidis (serogroup A), and
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OC NCBI_TaxID=65699, 491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
 RX MEDLINE=97252514; PubMed=9098083;
 RA Ropp P.A., Nicholas R.A.;
 RT "Cloning and characterization of the pona gene encoding penicillin-
 RT binding protein 1 from Neisseria gonorrhoeae and Neisseria
 RL meningitidis.";
 RL J. Bacteriol. 179:2783-2787(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
 RX MEDLINE=20225256; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
 RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RA "Complete DNA sequence of a serogroup A strain of Neisseria
 RL meningitidis Z2491.";
 RL Nature 404:502-506(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / Serogroup B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Cittiore H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,

RA Gill J., Scarlato V., Mesignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
RL MC58.";
RL Science 287:1809-1815(2000).
CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNITS) (BY SIMILARITY).
CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
CC similarity).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U80933; AAB52541.1; -;
DR EMBL; AL162753; CAB83943.1; -;
DR EMBL; AF002530; AAF42144.1; -;
DR TIGR; NWB1807; -;
DR InterPro; IPR001264; Transglycosyl.
DR InterPro; IPR001460; Transpeptidase.
DR Pfam; PF00912; Transglycosyl; 1.
DR Pfam; PF00905; Transpeptidase; 1.
DR ProDom; PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
KW Signal-anchor; Antibiotic resistance; Complete proteome.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 31 798 PERIPLASMIC (POTENTIAL).
FT DOMAIN 50 218 TRANSGLYCOSYLASE.
FT DOMAIN 378 700 TRANSEPTIDASE.
FT ACT_SITE 461 461 ACYLATED BY PENICILLIN (BY SIMILARITY).
SQ SEQUENCE 798 AA; 88147 MW; 0BAF4EF037A9977B CRC64;

Query Match 18.4%; Score 640; DB 1; Length 798;
Best Local Similarity 26.1%; Pred. No. 1.6e-29;
Matches 203; Conservative 121; Mismatches 272; Indels 182; Gaps 21;

QY 2 IYDNKNQLIADLSGRRVNAQANDPTDLVKAIVSTEDHRFDHGDIDIRILGAFRLNL 61
DB 52 IYDSAGEVIGMYGEORREPTKIGDFEVLNRNAVIAAEKRFYRHGVGVVARAAVGNV 111
QY 62 QSNLSQ-GGSTLTQOLIKLTSTSTSDOTISRKAQEAWLALQLEOKATQOELLTYINK 120
DB 112 VSGSVQSGASTITQOAKNFYLS---SEKTFTRKNEVLLAYKIEQSLSKILELYFNQ 168
QY 121 VYMSGNGYGMOTAAQNYGKDLNNLSLPOLLAGMPQAPNOYDPYSHPEAAQDRNLVL 180
DB 169 IYLGQRAYGFAAQAQIYFNKNVRDLTLAEAAAGLAPKAPSAYNPVTPNERAKLQKYL 228
QY 181 SEMKNGQYISAOYEKAVNTPTDGLQSLKASNPAYMDN-----YLKEVINO-VEEE 233
DB 229 NNMLEKMITVOORQALNEEL-----HYERFVRKIDQSALEYAEMVRQELYEK 277
QY 234 TGYNLLTTGDMVYTNVDQBAOK----- 255
DB 278 YGEDAYTGFKYTTVTRADHQKVAATEALRNDRGSSYRGAENYIDLKSEDEETV 337

QY 256 --HLWDIYNDEYV----- 267
DB 338 SOYLSGLYTVDKMVPVLDVTKKKNVVQLPGGRRVYTLDRRALGFAARAVNNEKMGEDR 397
QY 268 -----AYPDDELQVASTIVDYSNCKVIAQLCAQCARHSSNVSGFINQAVE 310
DB 398 IRGAVIRVKNNGRWAVVOEPLQGLVSLDAKTGAVALVGVGYDFHST---FNRAVQ 454
QY 311 TNRDWSGTMKPIITDYAPALEYGYVDSTATIVHDEPNYPCGN-----TPVYNWDRGYFG 364
DB 455 AMRQPGSTKPEFV-YSAALS-KGM--TASTVVDAPISLPKGPNGSVWTP-KNSDGRYSG 510
QY 365 NITLOVALQOQRNVPVETLNKVLNRAKTFNLGLIDYPSIIHYSNAISNTTESKKYK 424
DB 511 YITLQALTASKNWSIRILMSIGVGAQYIRRFGRSSSELPASISALGTGET----- 565
QY 425 ASSEKMAAAAFANGFTYKPMYIHKVVFSDGSEK-----FSNVGTFRAMKETAYMM 478
DB 566 -TPLKVAEAYSFANGYRVSSHVVDIKYDRDGLRLAQOMPLVAGONAPQAIIDPRNAYIM 624
QY 479 TDMKMTLVITYGGRN-AYLAWLPQAGKTGTSTNYTDEIEHNHKTSTQFVAPDELFAGYTRK 537
DB 625 YKIMQDVVRVGTARGAALGLRTDIAGKTGTN-----DN-----KDAWFVGFPND 669
QY 538 YSAVWVTGYSN-RLTPLVGNGLTAAKAVYKSMYTYLSEGSNPEDWNIPEGLY-RNGEYVF 595
DB 670 VTAVTYIGDFKPKSMGRVGVGGTIAVPVVDYMRFAKLGKQKGMKMGECVSSNGEYFM 729
QY 596 K-----NGARSTWNSPAQOPP-----STESSSSSSDSTSSSTSTPTNNST 639
DB 730 KERWVTDPLGLTLDNSGIAQPSSRAKDDGGAAGRGQAADDEVQDMQETPVLPSNT 787

RESULT 10
ID BBPA_NEIGO STANDARD; PRT; 797 AA.
AC 005131;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Penicillin-binding protein 1A (PBP-1a) (BBPa) [Includes: Penicillin-
DE insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TCase);
DE Penicillin-sensitive transpeptidase (EC 3.4.-) (DB-transpeptidase)].
GN MRCA OR PONA.
OS *Neisseria gonorrhoeae*.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; *Neisseria*.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-25.
RC STRAIN=FA19, AND ATCC 700825 / FA 1090;
RX MEDLINE=97252514; PubMed=9098083;
RA Ropp P.A., Nicholas R.A.;
RT "Cloning and characterization of the ponA gene encoding penicillin-
RT binding protein 1 from *Neisseria gonorrhoeae* and *Neisseria*
RT *meningitidis*.";
RL J. Bacteriol. 179:2783-2787(1997).
CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNITS) (BY SIMILARITY). ESSENTIAL FOR CELL WALL SYNTHESIS.
CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
CC similarity).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; 072876; AAB52536.1; -
 DR InterPro; IPR001264; Transglycosyl.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00912; Transglycosyl; 1
 DR Pfam; PF00905; Transpeptidase; 1.
 DR ProDom; PD001895; Transglycosyl; 1.
 KW Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
 KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
 KW Signal-anchor; Antibiotic resistance.
 FT INIT MET 0 0
 FT DOMAIN 1 8 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 9 29 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN) (POTENTIAL).
 FT DOMAIN 30 797 PERIPLASMIC (POTENTIAL).
 FT DOMAIN 49 217 TRANSGLYCOSYLASE.
 FT DOMAIN 377 699 TRANSEPTIDASE.
 FT ACT_SITE 460 460 ACYLATED BY PENICILLIN (BY SIMILARITY).
 SQ SEQUENCE 797 AA; 88364 MW; C7A0ID2BICAC9F3B CRC64;
 Query Match 18.2%; Score 635.5; DB 1; Length 797;
 Best Local Similarity 26.4%; Pred. No. 2.9e-29;
 Matches 207; Conservative 120; Mismatches 287; Indels 169; Gaps 22;
 QY 2 IYDNKQLIADGSRERRVNAQANDIPTDLVKAIVSIEDHRRFDHGRDITIRILGAFRLNL 61
 DB 51 IYSDAGEVIGMYGEORREFTKIGDPFVLNNAVIAAEKRFYRWGVDMGVAAVAGNV 110
 QY 62 QNSLQ-CGSTITQILKLTFTSTSDTISRKAQEWLAQLQEKATKQBIILYYINK 120
 DB 111 VSGVSQSGASTITQVANKFYLS----SEKTFTRKNEVLAYKIEQSLKSKILELYNQ 167
 QY 121 VYMSNGNTGMOTAYGKGLNNLSLPQALLAGMPOAPNOYDPYSHPEAAQDRRLVL 180
 DB 168 IYLGORAYGFASAAQIYFNKNVRLDLTAEAMLAGLPKAPSAYNTVNPRAKLQKYL 227
 QY 181 SEMKNGYISABQEKAVNTPI--TDGLOSLSKASNYPAYMDNYLKEVINQVEETGYNL 238
 DB 228 NNMLEEMKITVQORDQALNEELHYERFVRKIDQSALYAE-----VRRLEYKYGEDA 281
 QY 239 LTGMDVYTNVDQEAQ-----HLW 258
 DB 282 YTGKFKVITVTDRHOKAATEALRKALNFRDGRSSYGAENYIDLKSEDVEETVSQYLS 341
 QY 259 DTYNTDEYV-----AYPDE----- 273
 DB 342 GLYTVDKMVPVAVLDVTKKNVVIQPGRRVALDRALGFAARAVDNEKMGEDRIIRCA 401
 QY 274 -----LQVASTIVDSNGKVIQAGLARGHSSNVSGFINQAVETNRDW 315
 DB 402 VIRKNGNGRWAVQOEPLQGLVALVDKARTGAVALVGDFHSHKT---FNRAVQAMRQP 458
 QY 316 GSTMKPTIDYPALEYGVYDSTATVHDEPNYPCFN-----TPVYNWRGTYFCGNITLQ 369
 DB 459 GSTFAFV-YSAALSKGM--TASTVNDAPISLPKGGNGSVWTP-KNSDGRYSYITLR 514
 QY 370 VALQOSNRNPVAVETLUNKVGLNRAKFLNGLGIDYPSIHYSAISNTTESDKKYGASKE 429
 DB 515 QALTASKNWSIRILMSIGVGVAQYQYIRFGFRPSELPSALSMALGTGET-----TPLK 568
 QY 430 MAAAYAAAFANGTYTKYKPIIKHVFSDDSEK-----FSNVGTRAMKETATYMMTDMWK 483
 DB 569 VAEAYSVFANGGYRVSSSHVIDKIYDRDGLRAQMQPLVAGQNAQIDPRNAYIMYKIMQ 628
 QY 484 TVLTGTGRN-AYLAWLPQAKTGTSNYTDEIENHKTQSFPVDELFAGYTRYKYSNAV 542
 DB 629 DVVRVGTARGAAALGRDITAGKTGTN-----DN-----KDAWFGVGNPDVVTAV 673

QY 543 WTGYSN-RLTPVLGNGLTVAARKVYRSMMTYLSGSSNPEDWNIPEGLY-RNGEYVFNK--- 597
 DB 674 YIGFDKPKSMRGAGYGGTIAVPVWDYMRFAKLGKQKMGKPEGVVSSNGEYMKRWV 733
 QY 598 ---GARSTWNSAPQPPPTSSSSSSSTSTPSTNNSTTTNNNNNTQQSNWTP 654
 DB 734 TDPGLMLDNGSIAPQ--PSRRAKEDDAEVAENEQQRSDETRDQVQETP---VLPSTNDS 788
 QY 655 DQO 657
 DB 789 KOQ 791
 RESULT 11
 ID PBPA_NEICI STANDARD; PRT; 798 AA.
 AC 086088;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Penicillin-binding protein 1A (PBP-1a) (PBPla) [Includes: Penicillin-
 DE insensitive transglycosylase (EC 2.4.2.-) (Peptidoglycan TGase);
 DE Penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].
 GN MRCA OR PONA.
 OS Neisseria cinerea.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=483;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Ropp P.A., Nicholas R.A.;
 RT "Cloning and sequence analysis of the pona gene encoding penicillin
 RL binding protein 1 from Neisseria cinerea."
 CC Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
 CC PEPTIDOGYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
 CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
 CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
 CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
 CC SUBUNITS) (BY SIMILARITY).
 CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGYCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
 CC similarity).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC TRANSGLYCOSYLASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
 CC TRANSEPTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF085340; AAC34128.1; -
 DR InterPro; IPR001264; Transglycosyl.
 DR InterPro; IPR001460; Transpeptidase.
 DR Pfam; PF00912; Transglycosyl; 1.
 DR Pfam; PF00905; Transpeptidase; 1.
 DR ProDom; PD001895; Transglycosyl; 1.
 KW Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
 KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
 KW Signal-anchor; Antibiotic resistance.
 FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN) (POTENTIAL).
 FT DOMAIN 31 798 PERIPLASMIC (POTENTIAL).
 FT DOMAIN 50 218 TRANSGLYCOSYLASE.
 FT DOMAIN 414 700 TRANSEPTIDASE.
 FT ACT_SITE 461 461 ACYLATED BY PENICILLIN (BY SIMILARITY).
 SQ SEQUENCE 798 AA; 87843 MW; 5842ED4BCB9F06A CRC64;

```
Query Match      18.2%; Score 633; DB 1; Length 798;
Best Local Similarity 26.5%; Pred. No. 4.1e-29;
Matches 191; Conservative 115; Mismatches 248; Indels 166; Gaps 19;

QY 2 IYDNKNQIADLGSRRVNAQANDIPTDLVKAIVSIEDHRFFDHGIDTIRILGAFRLN 61
DB 52 IYSADEGEVIGIYGEORREFTKIGDFPEVLRNAVIAEDKRFYQHWGVVDVWGVARAVVGN 111
QY 62 QNSLQ-QGSTLTQOLIKLTFTSTSDQTSRKAQEAWLAIQLEKATKQKQKQKQKQK 120
DB 112 VAGGVQSGASTITQVAKNFYLS---SEKTRKNEALLAYKIQSLSKQKQKQKQK 168
QY 121 VYMSNGNYGMOTAAQNYGKOLNLSLPQALLAGMPQAPNOYDPSHPFAAQDRRLV 180
DB 169 IYLGORAYGFAAQQIYFNKQVRELTLAEVAMLAGLPRAPSAYNPVNPRAKQKQYL 228
QY 181 SEMKNQGYISAEQYKAVNTPTIDGLQSLKASNPAYMDN-----YLKEVINQ-VEEE 233
DB 229 NMMEERMITLQORQALNEEL-----HYERFVKIDQSAIYVAEMVRQELYEK 277
QY 234 TCYNLLTTGMVYVNVDOEAQK----- 255
DB 278 YGEDAYTOGFKYTYTTRDQKVAETALRKNLNRDGRSSYRGARSYIDLKSGDVEETV 337
QY 256 --HLMDIYNTDE----- 265
DB 338 SOYLSGLYTVDKWPAIVLDVTKRNVVQLPSGKRVTLDRSLGFAARAVNNEKMGESR 397
QY 266 -----YVAYPDDELQVASTVDVSGKVIQAIGARHQSNVSGFQINQAVE 310
DB 398 IRRGSVIRVRNNGRWVYVQEPQLQATLVSLDAKTCAVRALVGYDFHSEKTFNRAAQ 454
QY 311 TNRDNGSTMKPIITDAPALEYGVYDSTATIVHDEPNYP--GTNPVY---NWRDGYFCN 365
DB 455 ANRQGSFTKPEI-YSAALSKGM--TASTVNDADISLPKGANGSVWTPKNSDGRYSY 511
QY 366 IFLOYALQOSRNPVAVETLNKVLNRAKTFNLGLGIDYPSIHYNSAINTTESDKKYGA 425
DB 512 IFLQALTSKNNVSIIRILMSIGVGYAHEYIQRFQFKPSLPSMALGTGET----- 565
QY 426 SSEKAAAYAFANGCTYKPMYIHKVFSDCSEKE-----FSNVGTRFMKETTAYMT 479
DB 566 TPLKTAEAYSVFANGGYRVSSHVVDIKYSGDGLRQAQPLVAGQAPQADPRNAYIMY 625
QY 480 DMKQVTLVYGTGRN-AYLAWLPQACKTGTSNYTDEIEENHIKTSQFVAPDELFAGYTRY 538
DB 626 KIMQDVVRVGTARGAALGRSDIAGKTGTN-----DN-----KDAWFGVFNPDV 670
QY 539 SMAVWTGYSN-RLTPLVNGLTVAARVSRMMTYLSEGSNPDWNIPEGLY-RNGEFVFK 596
DB 671 VTAVYIGFDKPKSMGRAGVGGTIAVPVWVDMRFALKGGQGMKVPGEVVSNGEYMK 730

RESULT 12
ID BPBA_VIBCH STANDARD; PRT; 825 AA.
AC Q9KN05;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Penicillin-binding protein 1A (PBP1a) (PBP1a) [Includes: Penicillin-
DE insensitive transglycosylase (EC 2.4.2.-) (peptidoglycan TGase);
DE Penicillin-sensitive transpeptidase (EC 3.4.-.-) (DD-transpeptidase)].
GN MRCA OR PONA OR VC2635.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
```

```
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Olin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.K., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RL Nature 406:477-483(2000).
CC -!- FUNCTION: CELL WALL FORMATION. SYNTHESIS OF CROSS-LINKED
CC PEPTIDOGLYCAN FROM THE LIPID INTERMEDIATES. THE ENZYME HAS A
CC PENICILLIN-INSENSITIVE TRANSGLYCOSYLASE N-TERMINAL DOMAIN
CC (FORMATION OF LINEAR GLYCAN STRANDS) AND A PENICILLIN-SENSITIVE
CC TRANSEPTIDASE C-TERMINAL DOMAIN (CROSS-LINKING OF THE PEPTIDE
CC SUBUNIT) (BY SIMILARITY).
CC -!- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (By
CC similarity).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC TRANSGLYCOSYLASE FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE
CC TRANSEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004330; AAF95776.1; ALT_INIT.
DR TIGR; VC2635.
DR InterPro; IPR001264; Transglycosyl.
DR Pfam; PF00912; Transglycosyl; 1.
DR ProDom; PD001895; Transglycosyl; 1.
KW Peptidoglycan synthesis; Cell wall; Transferase; Glycosyltransferase;
KW Hydrolase; Multifunctional enzyme; Transmembrane; Inner membrane;
KW Signal-anchor; Antibiotic resistance; Complete proteome.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 7 27 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 28 825 PERIPLASMIC (POTENTIAL).
FT DOMAIN 48 216 TRANSGLYCOSYLASE.
FT DOMAIN 413 752 TRANSEPTIDASE.
FT ACT_SITE 471 471 ACYLATED BY PENICILLIN (BY SIMILARITY).
FT SEQUENCE 825 AA; 91986 MW; 8A9553DA2650B9D2 CRC64;
Query Match 17.7%; Score 617; DB 1; Length 825;
Best Local Similarity 24.3%; Pred. No. 3.6e-28;
Matches 200; Conservative 128; Mismatches 220; Indels 276; Gaps 28;

QY 1 KIYDNKNQIADLGSRRVNAQANDIPTDLVKAIVSIEDHRFFDHGIDTIRILGAFRLN 60
DB 49 QVFSQDGKLIQAQFGEKRRPRLKLEEMPKELIEAVIATEDSRVYEHYGFDPGIGTTRAAFAV 108
QY 61 LOSNLS-QGSTLTQOLIKLTFTSTSDQTSRKAQEAWLAIQLEKATKQKQKQKQK 119
DB 109 LASGASQSGASTITQOLARNFFLS---NEKVMYRKVEIFAIHIEQLLSKQKQKQK 165
QY 120 KYVMSNGNYGMOTAAQNYGKOLNLSLPQALLAGMPQAPNOYDPSHPFAAQDRRLV 179
DB 166 KIYLGSRVSGVAAQAQVFGKEVKDLTLGEIALIAGLPRAPSTMNPPIYSVERATNRNVV 225
QY 180 LSEKMQGYISAEQYKAVNTPTIDGLQSLKASNPAYMDNLYKEV-----INQVEET 234
DB 226 LQRMDEKVIYTKAEYDAARAEPVLPKYHCAETELNAP-----YVAEIRAWWVRYGEFA 280
QY 235 GYNLLTGMVYVNVDOE-----AQKHLMDI----- 260
DB 281 AY---TSGMNVYTVDSKLQRAANQAAINNLLAYDHRHGYRGAERKELMQVNPWSSQTL 337
```

THIS PAGE BLANK (USPTO)